

clone\_end:T7  
vector\_side:right  
misc\_feature  
144606..187852  
/note="assembly name:Contig41"  
SEQUENCE 55964 a 35742 c 36408 g 58936 t 802 others  
BEGIN

Query Match:  
Best Local Similarity 100.0%; Score 738; DB 2; Length 187852;  
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

244 AAGTTGGTCTCTCTTAACCTATCCATCCCTCCCTCAATTAAGTGAACACTCCAAAGT 323  
6477b AGGTTGGTCTCTCTTAACCTATCCATCCCTCCCTCAATTAAGTGAACACTCCAAAGT 64835  
424 GTTATTCCTTAAACACACTGATATCATTTAAAGCCACTAAGAGAGTGAATCAGAA 383  
6443b GTTATTCCTTAAACACACTGATATCATTTAAAGCCACTAAGAGAGTGAATCAGAA 64895  
384 TAAAGCAATCATATAGAAAGTATGATCAGACAGATCCACCCAGACCTTCCACCGTGA 443  
6489b TAAAGCAATCATATAGAAAGTATGATCAGACAGATCCACCCAGACCTTCCACCGTGA 64955  
444 ACACACACACACACACACACATAGACACACACACACACACACACACACACACACAC 503  
6495b ACACACACACACACACACACATAGACACACACACACACACACACACACACACACAC 65015  
504 GCAGGGGACACAGTATGACATCTCTGAGAGTGAGGAGCAGCTGGGAGGAGGAGGAGG 563  
6501b GCAGGGGACACAGTATGACATCTCTGAGAGTGAGGAGCAGCTGGGAGGAGGAGGAGG 65075  
564 AATATTCGGGGATTTCTGAGGGGACAGATGTTCTCCAGGGCTGATGGCAGTCTT 623  
6507b AATATTCGGGGATTTCTGAGGGGACAGATGTTCTCCAGGGCTGATGGCAGTCTT 65135  
624 CAGAAAGTACTGACTTCAAC 683  
6513b CAGAAAGTACTGACTTCAAC 65195  
684 AGGGCAGTCT 743  
6519b AGGGCAGTCT 65255  
744 GAGTGCAGGACCAAGGCTATGACAGGCTGATCTTCACTGGCAGCTTGTTCATGGCA 803  
6525b GAGTGCAGGACCAAGGCTATGACAGGCTGATCTTCACTGGCAGCTTGTTCATGGCA 65315  
804 CTTCAGAGTCACTCTCAATGTGGCTGATCCCTAAACCAAACTCTTAATTGGCCAGT 863  
6531b CTTCAGAGTCACTCTCAATGTGGCTGATCCCTAAACCAAACTCTTAATTGGCCAGT 65375  
864 ACCTGACACACAGACCAAGAGCTTCTACTGTGTGAGGCTCAACACATGCTCTGTG 923  
6537b ACCTGACACACAGACCAAGAGCTTCTACTGTGTGAGGCTCAACACATGCTCTGTG 65435  
924 CTGCACACTCACTGATATTTCTTAAGAAAGAAAGTCTTCTTAATTAAGAAAGAA 983  
6543b CTGCACACTCACTGATATTTCTTAAGAAAGAAAGTCTTCTTAATTAAGAAAGAA 65495  
984 AGAGCTCTTTGAAAGCT 1001  
6549b AGAGCTCTTTGAAAGCT 65513

158380/c AX458380 621 bp DNA linear PAT 08-JUL-2002  
US INITIATION Sequence 5 from Patent W00246411.  
ESSION AX458380  
STION AX458380.1 GI:21725048  
WORDS  
JRCSE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Lasek A.W., Krasnow R. and Baughn M.R.  
Intestinal proteins  
Patent: WO 0246411-A 5 13-JUN-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
Location/Qualifiers  
source  
1..621  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 142 a 138 c 179 g 162 t  
ORIGIN

Query Match:  
Best Local Similarity 99.8%; Score 558; DB 6; Length 621;  
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

210 TGTGTTATATATAGCTCTGATCATTTTATACAGTTTCCAGAACTCCCAAGATT 269  
621 TGTGTTATATATAGCTCTGATCATTTTATACAGTTTCCAGAACTCCCAAGATT 562  
270 GGTCTGCTAAATTAATTCATCTCTCTCAATTAAGTGAACAACTCCCAAGATTCTTAT 329  
561 GGTCTGCTAAATTAATTCATCTCTCTCAATTAAGTGAACAACTCCCAAGATTCTTAT 502  
330 TCCCTAAACACACACTGATATCATTTAAAGCAGTAAAGAGTGAAGTGAAGTGAAG 389  
501 TCCCTAAACACACACTGATATCATTTAAAGCAGTAAAGAGTGAAGTGAAGTGAAG 442  
390 AATCATATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 449  
441 AATCATATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 382  
450 CCACACGCGCCACACATATGACACAGCAGCACTCTCAACCTCTGAGATACCTAGCAGG 509  
381 CCACACGCGCCACACATATGACACAGCAGCACTCTCAACCTCTGAGATACCTAGCAGG 322  
510 GAGACAGTCACT 569  
321 GAGACAGTCACT 262  
570 GCGGGATTTCTGAGGGCAGAGATGTTCTTCCACAGCTGATGAGGATGTTCAAGAA 629  
261 GCGGGATTTCTGAGGGCAGAGATGTTCTTCCACAGCTGATGAGGATGTTCAAGAA 202  
630 AGTGTGACATTCACCAACAGGACACCTACTGTCAGCTTCACTCTCTTCTGGAGGGCC 689  
201 AGTGTGACATTCACCAACAGGACACCTACTGTCAGCTTCACTCTCTTCTGGAGGGCC 142  
690 AGTGTGCTCTGTTCTGCTGCTCATCCACCCAGTGAAGGGGTTGAGCTCTCTGAGATG 749  
141 AGTGTGCTCTGTTCTGCTGCTCATCCACCCAGTGAAGGGGTTGAGCTCTCTGAGATG 82  
750 CAAGAAACAGAGCTATGACAGGCTGATCTTCACTGAGGCAATTTGCAATGCACTTCCC 809  
81 CAAGAAACAGAGCTATGACAGGCTGATCTTCACTGAGGCAATTTGCAATGCACTTCCC 22  
810 AAGTCCACT 818  
21 AAGTCCACT 13

RESULT 10  
AK000134  
LOCUS AK000134 1704 bp mRNA linear PRI 22-FEB-2000  
DEFINITION Homo sapiens cDNA FLJ20127 fls, clone COL06176.  
ACCESSION AK000134  
VERSION AK000134.1 GI:7020022  
KEYWORDS  
oligo capping; fls (full insert sequence).  
SOURCE  
ORGANISM Homo sapiens

\_\_\_\_\_

799 TGGCACTTCCCAAGTCCACTTCGATGTGGC 829

Fri Nov 22 09:40:46 2002

us-09-729-454-3.nov1801i.rst

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:33:06 ; Search time 2189 Seconds

(without alignments)  
11956.102 Million cell updates/sec

Title: US-09-729-454-3

Perfect score: 1616

Sequence: 1 atccatgctaaagtaaac.....ataacaaaaaaaaa 1616

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: listing first 45 summaries

Database:

EST:  
1: em\_eatba:\*  
2: em\_eatbum:\*  
3: em\_estim:\*  
4: em\_estim:\*  
5: em\_estim:\*  
6: em\_estim:\*  
7: em\_estim:\*  
8: em\_hic:\*  
9: qb\_est1:\*  
10: qb\_est2:\*  
11: qb\_hic:\*  
12: qb\_est3:\*  
13: qb\_est4:\*  
14: qb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: qb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_juv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtc:\*  
22: em\_gss\_vtc:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

pred No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
1	268	16.6	358	9	AI833131	AI833131 at75d06.x
2	169	10.5	905	12	BG199761	BG199761 RST19055
3	148	9.2	343	12	BG201797	BG201797 RST21139
4	112	6.9	112	9	AI581835	AI581835 at94g08.x
5	111	6.9	946	12	BG217603	BG217603 RST37215
6	109	6.7	530	17	AO439960	AO439960 HS_5059_B

7	8	5.4	634	13	BI764226	BI764226 603045907
9	45	2.8	327	10	AW842215	AW842215 RCO-CN002
10	32	2.0	664	10	BB625007	BB625007 BB625007
11	32	2.0	367	11	AK019469	AK019469 Mus muscu
12	27	1.7	573	17	A2454451	A2454451 1M0256319
13	27	1.7	630	10	BB656245	BB656245 BB656245
14	24	1.5	378	12	BF721682	BF721682 mab21c06
15	24	1.5	446	10	BB623365	BB623365 BB623365
16	24	1.5	597	12	BC088025	BC088025 H3147F12
17	24	1.5	671	10	BB605016	BB605016 BB605016
18	23	1.4	227	13	RG42595	RG42595 a227b03.x
19	23	1.4	252	10	AW522484	AW522484 UI-R-B00-
20	23	1.4	279	10	AM087776	AM087776 UI-R-B00-
21	23	1.4	357	9	AI136893	AI136893 UI-R-C2P-
22	23	1.4	375	12	BO514625	BO514625 dae09f12
23	23	1.4	403	9	AI482526	AI482526 vq49d02.x
24	23	1.4	418	10	AW494429	AW494429 UI-M-BH3-
25	23	1.4	434	13	BI494352	BI494352 f020a06.x
26	23	1.4	436	9	AL047005	AL047005 df109e02
27	23	1.4	472	10	AW435156	AW435156 UI-R-B00P
28	23	1.4	490	13	BM445719	BM445719 UI-H-DH1
29	23	1.4	492	14	BO018316	BO018316 UI-H-DH1
30	23	1.4	499	12	CG3956	CG3956 UI-R-B00P
31	23	1.4	521	12	BS79297	BS79297 dab85b07
32	23	1.4	536	13	BF510845	BF510845 UI-R-B1-
33	23	1.4	537	13	BM385280	BM385280 UI-R-B1-
34	23	1.4	544	14	BM989987	BM989987 UI-R-B1-
35	23	1.4	554	10	AW526128	AW526128 UI-H-D10-
36	23	1.4	564	13	BI402698	BI402698 UI-R-B01-
37	23	1.4	580	13	BM030140	BM030140 UI-R-B01-
38	23	1.4	588	14	BM973554	BM973554 UI-R-B01-
39	23	1.4	598	17	BM323298	BM323298 UI-CF-EC1
40	23	1.4	620	14	BO574501	BO574501 UI-H-E21-
41	23	1.4	622	9	AU039013	AU039013 UI-H-E21-
42	23	1.4	622	14	BM683810	BM683810 UI-E-E11-
43	23	1.4	623	14	BM971486	BM971486 UI-CF-EC1
44	23	1.4	639	14	BO190688	BO190688 UI-R-DN1-
45	23	1.4	639	14	BO190688	BO190688 UI-R-DN1-

ALIGNMENTS

RESULT 1  
LOCUS AI833131/c  
DEFINITION at75d06.x1 Barstead colon HPRB7 Homo sapiens cDNA clone  
IMAGE:2377835.3', similar to SW:BB61-RABIT 005004 BRUSH BORDER 61.9  
ID: PROTEIN PRECURSOR. ; mRNA sequence.  
ACCESSION AI833131  
VERSION AI833131  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 358)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,  
Kritman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,  
J., Moore, B., Scheinberg, R., Stepien, M., Tan, F., Theising, B.,  
White, Y., Wylie, T., Waterston, R., and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@imgl.lnl.gov) for further information.  
Seq primer: -40UP from Gdbco.

TITLE  
JOURNAL  
COMMENT

```

FEATURES
    location/Qualifiers
    1. 358
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:237835"
    /clone_lib="Barstead colon HPLRB7"
    /sex="male"
    /dev_stage="adult, age 25"
    /lab_host="DH10B (phage resistant)"
    /note="Organ: colon; Vector: pTR73b-Pac (Pharmacia) with a
    modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st
    strand cDNA was primed with a Not I - oligo(dT) primer [5'
    TTTAGACATCTGACGTGAGGAGCGCCCGCCCTTTTGTGTTTTTTTTTTT
    3']; double-stranded cDNA was ligated to Eco RI adaptors
    [5' ACTTACACTAGTAT 3' and 5' ATTCTAGTGG 3'], digested
    with Not I and cloned into the Not I and Eco RI sites of
    the modified pTR73 vector. Library constructed by Bob
    Barstead."

BASE COUNT      106 a      60 c      74 g      118 t

ORIGIN
Query Match      16 6%; Score 268; DB 9; Length 358;
Best Local Similarity 100.0%; Pred. No. 1.3e-76;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1328 AACCCCTTACGACCTCTTCTTGAGAAAGCCGACACATAGTATCATCAAAACGAA 1387
1329 TTTCTGACATCTGACATCTCTTCTTGAGAAAGCCGACACATAGTATCATCAAAACGAA 1387
1330 TTTCTGACATCTGACATCTCTTCTTGAGAAAGCCGACACATAGTATCATCAAAACGAA 1387
268 AACGCATTCACGATCTTCTTGAGAAAGCCGACACATAGTATCATCAAAACGAA 209
269 TTTCTGACATCTGACATCTCTTCTTGAGAAAGCCGACACATAGTATCATCAAAACGAA 209
270 TTTCTGACATCTGACATCTCTTCTTGAGAAAGCCGACACATAGTATCATCAAAACGAA 209
1388 AACATCGGAGAGATATCAATGATGAGAGAAAGATTAGTACTTCATGTTACATTCOA 1447
1389 AACATCGGAGAGATATCAATGATGAGAGAAAGATTAGTACTTCATGTTACATTCOA 1447
208 AACATCGGAGAGATATCAATGATGAGAGAAAGATTAGTACTTCATGTTACATTCOA 149
209 AACATCGGAGAGATATCAATGATGAGAGAAAGATTAGTACTTCATGTTACATTCOA 149
210 AACATCGGAGAGATATCAATGATGAGAGAAAGATTAGTACTTCATGTTACATTCOA 149
1448 TATCTCATCTCAATAAGACATTTTCCAGAGATCTGAGTAGATCATCTGATGCCGGGAT 1507
1449 TATCTCATCTCAATAAGACATTTTCCAGAGATCTGAGTAGATCATCTGATGCCGGGAT 1507
211 TATCTCATCTCAATAAGACATTTTCCAGAGATCTGAGTAGATCATCTGATGCCGGGAT 1507
Db 148 TATCTCATCTCAATAAGACATTTTCCAGAGATCTGAGTAGATCATCTGATGCCGGGAT 89
1508 ATACCAATTGCATATGCGCAAAATAATATGACACCACCTCAACATGTATGCGGAAATAG 1567
1509 ATACCAATTGCATATGCGCAAAATAATATGACACCACCTCAACATGTATGCGGAAATAG 1567
1510 ATACCAATTGCATATGCGCAAAATAATATGACACCACCTCAACATGTATGCGGAAATAG 1567
Db 88 ATACCAATTGCATATGCGCAAAATAATATGACACCACCTCAACATGTATGCGGAAATAG 29
2Y 1568 ATTAAATATTATTAACTAATATTTGTT 1595
2Y 1569 ATTAAATATTATTAACTAATATTTGTT 1595
2Y 1570 ATTAAATATTATTAACTAATATTTGTT 1595
Db 28 ATTAAATATTATTAACTAATATTTGTT 1

RESULT 2
BG199761      905 bp      mRNA      linear      EST 21-APR-2001
LOCUS        BG199761
DEFINITION   RST19055 Albersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION    BG199761
VERSION      BG199761.1 GI:13721448
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 905)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J
., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
., E., Veloso,N., Kilka,A., Hess,J., Colhen,K., Lo,K., Offenbacher
., J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
Contact: Scott J. Cain
Albersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@albersys.com

```

FEATURES		High quality sequence stop: 387.	
Location/Qualifiers			
source	1..905		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="Athersys RAGE Library"		
	/cell_line="HT1080"		
	/note="See 'Creation of genome-wide Protein Expression Libraries using Random Activation of Gene Expression' Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."		
BASE COUNT	228 a 226 c 296 g 154 t		
ORIGIN			
Query Match	10.5% Score 169; DB 12; Length 905;		
Best Local Similarity	99.1% Pred. No. 6e-45;		
Matches 319; Conservative	0; Mismatches 3; Indels 0; Gaps 0		
Db	352 ATTAAGCCACTAACAGAGACTGAACTGAGATTAAGCAAAATCATAGAGAAATCTACATCA 411		
Qy			
Db	73 ATTAAGCCACTTAACAGAGACTGAACTGAGATTAAGCAAAATCATAGAGAAATCTACATCA 132		
Qy	412 GCAGATCCGACCCAGACCTTCCACGAGTAAACACACACACACACACATACAC 471		
Db	133 GCAGATCCGACCCAGACCTTCCACGAGTAAACACACACACACACACATACAC 192		
Qy	472 AGCCACCATCTCTCAACCTTGAGATTAAGTACTGACAGGGAGACACACATCTCTCTCT 531		
Db	193 AGCCACCATCTCTCAACCTTGAGATTAAGTACTGACAGGGAGACACACATCTCTCTCT 252		
Qy	532 GGAGGTGAGGAGACCACTTGGAGCGCGAGAACCAATATGGCGGGATTTCTCTAGGGCCAC 591		
Db	253 GGAGGTGAGGAGACCACTTGGAGCGCGAGAACCAATATGGCGGGATTTCTCTAGGGCCAC 312		
Qy	592 GATGTTCTCCCGAGCGGTGATGGACAGTCTCTGAGGAAGTACTGATTCACCAACGCG 651		
Db	313 GATGTTCTCCCGAGCGGTGATGGAGAGTCTCTGAGGAAGTACTGATTCACCAACGCG 372		
Qy	652 CACCTACCTGTCACTTCACT 673		
Db	373 CACCTACCTGTCACTTCACT 394		
RESULT 3			
LOCUS	BG201797 343 bp mRNA linear EST 21-APR-2001		
DEFINITION	RST21139 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.		
ACCESSION	BG201797		
VERSION	BG201797.1 GI:13723484		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 343)		
AUTHORS	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,		
	Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J.		
	, Lener,L., Costanzo,D., McElligott,K., Boozier,S., Mays,R., Smith,		
	,E., Veloso,N., Kilka,A., Hess,J., Cohren,K., Lo,K., Olenbacher		
	,J., Danzig,J. and Ducar,M.		
TITLE	Creation of genome-wide protein expression libraries using random		
	activation of gene expression		
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)		
MEDLINE	21227151		
COMMENT	Contact: Scott J. Cain		
	Athersys, Inc.		
	3201 Carnegie Ave, Cleveland, OH 44115, USA		
	Tel: 216 431 9900		
	Fax: 216 361 9596		
	Email: scain@atersys.com		
	High quality sequence stop: 293.		

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:33:06 : Search time 2189 Seconds  
(without alignments)  
11956.102 Million cell updates/sec

Title: US-09-729-454-3

Perfect score: 1616  
Sequence: 1 atccatgcataaagtaaca.....aatacaaaaaaaaaaaaaa 1616

Scoring table: ONTC\_NTC  
Gapop 60 0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

EST.\*  
1: em\_estbda.\*  
2: em\_estbhm.\*  
3: em\_estln.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hnc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hnc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estcom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_liv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vit.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rtd.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	268	16.6	358 9	AI833131 at75d06.x
2	169	10.5	905 12	BG199761 RST19055
3	148	9.2	343 12	BG201797 RST11339
4	112	6.9	112 9	AI581835 at94908.x
5	111	6.9	946 12	BG217603 RST17315
6	109	6.7	530 17	AQ439960 HS_5059_B

Result 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AI833131/c	AI833131	358 bp mRNA	EST 13-JUL-1999			human	Homo sapiens	1 (bases 1 to 358)				
	at75d06.x1	Barstead colon HPLRB7 Homo sapiens cDNA clone				human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	IMAGE:2377835 3'	similar to SW:BB61.PABIT Q05004 BRUSH BORDER 61.9				human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	KD PROTEIN PRECURSOR.	;', mRNA sequence.				human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	AI833131					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	AI833131.1	GI:5455111				human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	WashU-NCI human EST Project					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Unpublished (1997)					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Contact: Wilson RK					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Washington University School of Medicine					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Tel.: 314 286 1800					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Fax: 314 286 1810					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Email: estewatson.wustl.edu					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	This clone is available royalty-free through LNL; contact the					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	IMAGE Consortium (info@image.llnl.gov) for further information.					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
	Seq primer: -40up from Gibco.					human	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					

## ALIGNMENTS

RESULT 1  
LOCUS AI833131/c  
DEFINITION at75d06.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone  
IMAGE:2377835 3' similar to SW:BB61.PABIT Q05004 BRUSH BORDER 61.9  
KD PROTEIN PRECURSOR. ;', mRNA sequence.  
ACCESSION AI833131  
VERSION AI833131  
KEYWORDS EST.  
SOURCE human  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 358)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getse, G., Joat, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Thelning, B., White, Y., Wyllie, T., Watson, R., and Wilson, R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel.: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40up from Gibco.





LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
B1764226	634 bp	MRNA	linear	EST 25-SEP-2001	
603045907P1	NIH_MGC_116	Homo sapiens	cDNA clone IMAGE:5186472.5,		
B1764226					
B1764226.1	GI:15755804				
EST.					
human.					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eulhelia; Primates; Catarrhini; Homiidae; Homo.					
NIH-MGC	http://mgc.nci.nih.gov/.				
1 (bases 1 to 634)					
Unpublished (1999)					
Contact: Robert Strassburg, Ph.D.					
Email: ccrabb@remail.nih.gov					
Tissue Procurement: Life Technologies, Inc.					
cDNA Library Preparation: Life Technologies, Inc.					
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Incyte Genomics, Inc.					
Clone distribution: MGC clone distribution information can be					
found through the I.M.A.G.E. Consortium/LLNL at:					
http://image.llnl.gov					
plate: L1AM11465 row: m column: 01					
High quality sequence stop: 634.					
Location/Qualifiers					
1..634					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/clone="IMAGE:5186472"					
/clone_1lb="NIH_MGC_116"					
/lab_host="DH10B"					
/note="Organ: pooled colon, kidney, stomach; Vector: PCWV-SPORT6; Site:1; Nof: Site:2; EORY (destroyed); RNA source anonymous pool of 3 colons; age 26 yo male; 49 yo female; 71 yo male colon; 46 yo male kidney; and pool of 2 stomachs; 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (PCWV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and lnced for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC library."					
BASE COUNT	180 a	175 c	140 g	139 t	
ORIGIN					
Query Match	5.4%	Score 88:	DB 13:	Length 634:	
Best Local Similarity	100.00:	Pred. No. 9.86-19:			
Matches 88:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:	
OY	351	CATTAAAGCCACTACAGAGACTGACTGAGATTAAGAAATATGACAACTGATC	410		
	175	CATTAAAGCCACTACAGAGACTGACTGAGATTAAGAAATATGACAACTGATC	234		
OY	411	AGCAGATCCACCCACAGACTTTCACCA	438		
	235	AGCAGATCCACCCACAGACTTTCACCA	262		
Db					
RESULT 8					
LOCUS	AM842215/c				
DEFINITION	AM842215	327 bp	MRNA	linear	EST 18-MAY-2000
ACCESSION	AM842215				
VERSION	AM842215.1				
KEYWORDS	GI:7936198				
EST.					
SOURCE	human.				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eulhelia; Primates; Catarrhini; Homiidae; Homo.					





LOCUS	AC019469	367 bp	MRNA	linear	HTC 19-Jan-2002
DEFINITION	Mus musculus 14 days embryo liver cDNA, RIKEN full-length enriched library, clone:4432416J03; homolog to BRUSH BORDER 61.9 KDA PROTEIN PRECURSOR, full insert sequence.				
ACCESSION	AC019469				
VERSION	AC019469.1	GI:12859693			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) 14 days embryo liver cDNA to mRNA, clone_11b;RIKEN full-length enriched mouse cDNA library clone:4432416J03.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.				
REFERENCE	1	Garnincel, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2	Garnincel, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
TITLE	Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042195				
REFERENCE	3	Shibata, K., Itoh, M., Alizawa, K., Nagaoaka, S., Sasaki, N., Carnincel, P., Konno, H., Akiyama, J., Nishi, K., Katsuna, T., Tashiro, H., Itoh, M., Sui, N., Ishii, Y., Nakamura, S., Hazama, M., Nishize, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Maehariki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawaji, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harai, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Alizawa, K., Itawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gotobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissel, C., King, B., Kochwa, H., Kuehn, P., Lewis, S., Matsuo, Y., Nakado, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Carninci, L. R., Barsh, G., Blake, K., Boffelli, D., Bojunga, A., Cantarel, C., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Togo, O., K., Wang, K. H., Welter, C., Whitlaker, C., Wilmink, J., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.			
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085660				
PUBMED	11217851				
REFERENCE	5 (Pages 1 to 3267)				
AUTHORS	Adachi, J., Alizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, Y., Hata, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F.,				

TITLE	Direct Submission			
JOURNAL	Submitted (18-ANG-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (1677-1), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: yoshihide-res@sc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
COMMENT	Please visit our web site ( <a href="http://genome-gsc.riken.go.jp/">http://genome-gsc.riken.go.jp/</a> ) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer: 5' GAGGAGACAGATCCACAGAGCTCTTTTCTTTTCTTTT 3'. cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rct = 10.0 and subclaction to Rct = 100.0. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGGAGAGGATCTCCAGTATTATTAATTATCCGCCCCGCCGCC 3'. cDNA was cleaved with BamHI and XhoI. cDNA of size compressed between 3 and 7 kb was selected before cloning. Vector: a modified phluescript KS(+), attested bulk excision from lambda FLC 1. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B			
FEATURES	Location/Qualifiers			
SOURCE	1..3267			
	/organism="Mus musculus"			
	/strain="C57BL/6j"			
	/db_xref="GenBank:DB:4432416303"			
	/db_xref="MGD:MG1:1912173"			
	/db_xref="taxon:10090"			
	/clone="4432416303"			
	/tissue_type="liver"			
	/clone_id="RIKEN full-length enriched mouse cDNA library"			
	/dev_stage="14 days embryo"			
	205..1881			
	/note="data source: SRR, source key: G55004, evidence: 188 homolog to data BRUSH BORDER 61.9 KDA PROTEIN PRECURSOR putative"			
	/codon_start=1			
	/protein_id="BA831740.1"			
	/db_xref="GI:12856594"			
	/db_xref="MGD:MG1:1925502"			
	/translation="MRRLSPRLILSLSPNLSAOKFLVLIIFAEVWFTSKDHTT FVHLNRIILRRMSIFKEFLHSEELKTPASVLAELATALELNGQIPPEPVIT SSTSAQSTGATIHNPQRTYQVCDQNLNVLIAAKYQFNKEGQFLRARSSTPAKRSQV GASKVDFNNGTYLISFTLMEGPVSLILMPSSEGSALMRHNGYKLIIFALAGST FLNISTVLEEGGLTISAELOCYDADRHAEKCTKJQPEALHTHTSNNSITF YLSIEELKFLPRNTGVEVYVNSLIVSFCNKKKTKKNGKQCGEETISGATLKSH WITHCONFEPAIDINDINCCTKRLIYLMGSKTQGLYLPKRYKTLKFPDHNAGG KFTHLIDTBEHLYPMQKKSHRPVTLNFSKSDNTFHEDVAGHSGTALVSPHSC QHEPFPINFTIRALINKNAIFERLNSPEKTYITENIEHIEHVFIPDHPFOS ONLIIRIFPELDNGVIDAMQWTVAVRSBDVHPESVLSQITPMTN1C"			
	3250..3255			
	/note="putative"			
	3267			
	/note="putative"			
BASE COUNT	954 a 726 c 673 g 914 t			
ORIGIN	Query Match			
	2.08; Score 32; DB 11; Length 3267;			

Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 716 CACCCAGTGAAGCGCTCTCAGCTCTGGAG 747  
Db 784 CACCCAGTGAAGCGCTCTCAGCTCTGGAG 815

MS011 11  
A2454451

LOCUS 573 bp DNA linear GSS 04-OCT-2000  
DEFINITION IM0256J19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

ACCESSION clone UUGC1M0256J19 F, DNA sequence.

VERSION A2454451

KEYWORDS A2454451.1 GI:10612576

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah  
Bldg. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0256 row: J column: 19  
Seq primer: CGTGTAAACGACGCCACGT  
Class: plasmid ends  
High quality sequence stop: 573.

FEATURES

Source

Location/Qualifiers  
1..573  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0256J19"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. coli strata XL10-Gold, TI-resistant, F-"  
/note="Vector: pMD42ny; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pMD42 (g1473211419b1AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

Best Local Similarity 100.0%; Pred. No. 48;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 662 GTTCAGCTTCACCTCTTCTGGAGAGGC 688  
Db 511 GTTCAGCTTCACCTCTTCTGGAGAGGC 537

RS01T 12  
BB656245

LOCUS 630 bp mRNA linear EST 26-OCT-2001  
DEFINITION BB656245 RIKEN full-length enriched, 12 days embryo spinal ganglion

ACCESSION BB656245 cdna clone D130007D15 5', mRNA sequence.

VERSION BB656245.1 GI:16490073

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

RIKEN Mouse ESTs (Arikawa, T., et al. 2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for genome exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gscc.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
Wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Matsuhira, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, T., Kita, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
Y., and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamazaki, I., Aizawa  
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (http://genome-gsc.riken.go.jp) for  
further details.  
e mouse tissues.

FEATURES

Source

Location/Qualifiers  
1..630  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="D130007D15"  
/clone\_lib="RIKEN full-length enriched, 12 days embryo  
spinal ganglion"  
/tissue\_type="spinal ganglion"  
/dev\_stage="12 days embryo"

Query Match

1.7%; Score 27; DB 17; Length 573;



```

/clone="G430027D22"
/clone-lib="RIKEN full-length enriched, pooled cell lines"
/notes="pooled cell lines : (cell_line-CRL-1751 WEHI 164),
(cell_line-CRL-2116 JC), (cell_line-RCB-0035 WEHI-3),
(cell_line-RCB-0464 Meth-A), (cell_line-RCB-0545 OHRA),
(cell_line-RCB-0559 K-1. Fl), (cell_line-RCB-1283 B16
melanoma), (cell_type-B cells, cell_line-CRL-1702 WEHI 231
), (cell_type-B lymphoid cells, cell_line-CRL-2065 M17C-1),
(tissue_type-B lymphoid stem cell, cell_line-CRL-2070 NE),
(tissue_type-bone marrow, cell_line-RCB-0544 MBT-2),
(cell_line-CRL-2028 SR-4987), (tissue_type-colon,
cell_line-RCB-0549 CLe-H3), (tissue_type-kidney,
cell_line-CCL-142 PAC), (tissue_type-sundundibular gland,
cell_line-CRL-1734 SCA-9, clone 15), (strain-BALB/c,
cell_type-B cells, cell_line-CRL-1669 BCL1 clone 13.20-3B3
), (strain-C3H, tissue_type-brain, cell_line-CRL-1443
HC3H1)"

```

BASH COUNT 110 a 134 c 110 g 92 t

ORIGIN

Query Match 1.5%; Score 24; DB 10; Length 446;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 CAGCAGATCCACCCAGACCTTTC 433  
|||||

DB 12 CAGCAGATCCACCCAGACCTTTC 35

RESULT 15  
LOCUS H008025  
DEFINITION H3147F12-5 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone  
H3147F12 5', mRNA sequence.  
ACCESSION H008025  
VERSION H008025.1 GI:12570589  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 597)  
AUTHORS Kargul,G.D., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka  
T.S., Carter,M.G. and Ko,M.S.H.  
TITLE Verification and initial annotation of NTA mouse 15K cDNA clone set  
JOURNAL Unpublished (2001)  
COMMENT Other ESTs: H3147P12-3  
Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
This clone set has been freely distributed to the community. Please  
visit http://igsun.grc.nia.nih.gov/cDNA/15K.html for details.  
Plate: H3147 row: F column: 12  
Seq primer: -21M13 Reverse  
High quality sequence stop: 597  
POLY-A-NO

FEATURES  
SOURCE location/Qualifiers  
1..597  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taeST:H3147P12-5"  
/db\_xref="taxon:10090"  
/clone="H3147P12"  
/clone-lib="NTA Mouse 15K cDNA Clone Set"  
/sex="Clones arrayed from a variety of cDNA libraries"  
/dev\_stage="Clones arrayed from a variety of cDNA  
libraries"  
/lab\_host="DH10B"  
/notes="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; This  
clone is among a rearranged set of 15,247 clones from 11

embryo cDNA libraries (including preimplantation stage  
embryos from unfertilized egg to blastocyst; embryonic  
part of E7.5 embryos, extraembryonic part of E7.5 embryos  
, and E12.5 female mesonephros/gonad) and one newborn  
ovary cDNA library. Average insert size 1.5 kb. All  
source libraries are cloned unidirectionally with Oligo(dT  
)-Not primers. References include: (1) Genome-wide  
expression profiling of mid-gestation placenta and embryo  
using a 15,000 mouse developmental cDNA microarray, 2000,  
Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)  
Large-scale cDNA analysis reveals phased gene expression  
patterns during preimplantation mouse development, 2000,  
Development, 127: 1737-1749; (3) Genome-wide mapping of  
unselected transcripts from extraembryonic tissue of  
7.5-day mouse embryos reveals enrichment in the t-complex  
and under representation on the X chromosome, 1998, Hum  
Mol Genet 7: 1967-1978."

BASH COUNT 169 a 151 c 138 g 139 t

ORIGIN

Query Match 1.5%; Score 24; DB 12; Length 597;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 CAGCAGATCCACCCAGACCTTTC 433  
|||||

DB 449 CAGCAGATCCACCCAGACCTTTC 472

Search completed: November 22, 2002, 04:29:06  
Job time : 2198 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

UM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:38:36 ; Search time 866 seconds  
(without alignments)  
706,730 Million cell updates/sec

Title: US-09-729-454-3  
Perfect score: 1616  
Sequence: 1 aatcctatgctaaagtaaacata.....aataacaaaaaaaaaaaaa 1616

Scoring table: OLIGO\_MUC  
Gapop 60.0, Gapext 60.0

Searched: 335578 seqs, 169365133 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database:

Published Applications, NA.\*  
1: /cgn2\_6/p/odata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/p/odata/1/pubpna/PC1\_NEM\_PUB.seq.\*  
3: /cgn2\_6/p/odata/1/pubpna/US06\_NEM\_PUB.seq.\*  
4: /cgn2\_6/p/odata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/p/odata/1/pubpna/US07\_NEM\_PUB.seq.\*  
6: /cgn2\_6/p/odata/1/pubpna/PC1US\_PUBCOMB.seq.\*  
7: /cgn2\_6/p/odata/1/pubpna/US08\_NEM\_PUB.seq.\*  
8: /cgn2\_6/p/odata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/p/odata/1/pubpna/US09\_NEM\_PUB.seq.\*  
10: /cgn2\_6/p/odata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/p/odata/1/pubpna/US10\_NEM\_PUB.seq.\*  
12: /cgn2\_6/p/odata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/p/odata/1/pubpna/US60\_NEM\_PUB.seq.\*  
14: /cgn2\_6/p/odata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1616	100.0	1616	10	US-09-729-454-3
2	1005	62.2	1005	10	US-09-729-454-30
3	560	34.7	1212	9	US-09-981-353-144
4	558	34.5	621	10	US-09-729-454-5
5	488	30.2	841	9	US-09-981-353-145
6	409	25.3	545	10	US-09-729-454-6
7	375	23.2	394	10	US-09-729-454-9
8	240	14.9	240	10	US-09-729-454-4
9	213	13.2	236	10	US-09-729-454-7
10	198	12.3	554	9	US-10-046-935-181
11	93	5.8	580	9	US-10-046-935-388
12	88	5.4	661	10	US-09-729-454-11
13	88	5.4	1545	10	US-09-729-454-31
14	78	5.4	2248	10	US-09-729-454-10
15	73	4.5	414	10	US-09-729-454-8
16	47	2.9	518	10	US-09-729-454-12
17	34	2.1	244	10	US-09-729-454-18
18	34	2.1	270	10	US-09-729-454-19
19	34	2.1	291	10	US-09-729-454-17

20	32	2.0	206	10	US-09-729-454-16	Sequence 16, Appl
21	32	2.0	257	10	US-09-729-454-22	Sequence 22, Appl
22	25	1.5	288	10	US-09-729-454-20	Sequence 20, Appl
23	24	1.5	275	10	US-09-729-454-21	Sequence 21, Appl
24	24	1.5	394	10	US-09-729-454-9	Sequence 9, Appl
25	22	1.4	91	10	US-09-876-574-5331	Sequence 5331, Ap
26	22	1.4	144	10	US-09-729-454-27	Sequence 27, Appl
27	22	1.4	250	10	US-09-729-454-24	Sequence 24, Appl
28	22	1.4	357	10	US-09-960-352-8389	Sequence 8389, Ap
29	22	1.4	396	10	US-09-960-352-8389	Sequence 40, Appl
30	22	1.4	4158	10	US-09-920-300A-507	Sequence 34, Appl
31	21	1.3	337	10	US-09-833-381-1895	Sequence 507, App
32	21	1.3	337	12	US-10-033-558-507	Sequence 507, App
33	21	1.3	365	10	US-09-783-580-2675	Sequence 2675, Ap
34	21	1.3	380	9	US-10-079-623-188	Sequence 36, App
35	21	1.3	381	9	US-10-079-623-36	Sequence 13, Appl
36	21	1.3	462	10	US-09-729-454-13	Sequence 1895, Ap
37	21	1.3	531	10	US-09-729-454-14	Sequence 28, Appl
38	21	1.3	802	10	US-09-833-381-1895	Sequence 23, Appl
39	21	1.3	2456	10	US-09-965-313-3	Sequence 750, App
40	20	1.2	262	10	US-09-729-454-28	Sequence 581, App
41	20	1.2	276	10	US-09-729-454-23	Sequence 153, App
42	20	1.2	625	10	US-09-770-149-730	Sequence 23, Appl
43	20	1.2	864	10	US-09-770-445-581	
44	20	1.2	1448	10	US-09-764-853-153	
45	20	1.2	1967	10	US-09-887-586A-23	

#### ALIGNMENTS

RESULT 1  
US-09-729-454-3  
Sequence 3, Application US/09729454  
Patent No. US20020137038A1  
GENERAL INFORMATION:  
APPLICANT: Lasek, Amy W.  
APPLICANT: Krasnow, Randi E.  
TITLE OF INVENTION: INTESTINAL PROTEINS  
FILE REFERENCE: PC-0028 CIP  
CURRENT APPLICATION NUMBER: US/09/729, 454  
CURRENT FILING DATE: 2000-12-04  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PERL Program  
SEQ ID NO 3  
LENGTH: 1616  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US20020137038A1 3229449CB1  
US-09-729-454-3

Query Match 100.0% Score 1616; DB 10; Length 1616;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 ATCCATGCTAAAGGTAACAACTGCACTTATATTCGCAATTTATTTGGTATAGCA 60  
OY 1 ATCCATGCTAAAGGTAACAACTGCACTTATATTCGCAATTTATTTGGTATAGCA 60  
DB 1 ATCCATGCTAAAGGTAACAACTGCACTTATATTCGCAATTTATTTGGTATAGCA 60  
OY 1 ATCCATGCTAAAGGTAACAACTGCACTTATATTCGCAATTTATTTGGTATAGCA 60  
DB 61 GAGGTATGCGAGTACGACACTGCTGCTGAGAGAAATTCGCACTGACCTGCGCGAG 120  
OY 61 GAGGTATGCGAGTACGACACTGCTGCTGAGAGAAATTCGCACTGACCTGCGCGAG 120  
DB 61 GAGGTATGCGAGTACGACACTGCTGCTGAGAGAAATTCGCACTGACCTGCGCGAG 120  
OY 61 GAGGTATGCGAGTACGACACTGCTGCTGAGAGAAATTCGCACTGACCTGCGCGAG 120  
DB 121 AGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
OY 121 AGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 121 AGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
OY 121 AGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
DB 181 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
OY 181 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
DB 181 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240  
OY 181 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240

```

OY 241 TACAGTTTCCAGAACTCCACAAAGGTTGGTCTGCTCTAACTATCCATCTCCCTCCA 300
DB 241 TACAGTTTCCAGAACTCCACAAAGGTTGGTCTGCTCTAACTATCCATCTCCCTCCA 300
OY 301 TTACTGGAAACACTCCACAAAGTCTTATTCCTTAAACACCACTAGTATTCATTAAGCC 360
DB 301 TTACTGGAAACACTCCACAAAGTCTTATTCCTTAAACACCACTAGTATTCATTAAGCC 360
OY 361 ACTACAGAGACTGAACTCAGATTAAGAAATCTATAGAAATCTGATCAGCAGATCCC 420
DB 361 ACTACAGAGACTGAACTCAGATTAAGAAATCTATAGAAATCTGATCAGCAGATCCC 420
OY 421 ACCGAGACTTTCCACCCAGCGTGAACACACACAGCGCCACACATAGCAGACCAT 480
DB 421 ACCGAGACTTTCCACCCAGCGTGAACACACAGCGCCACACATAGCAGACCAT 480
OY 481 CCTCAACCTCGAGATAGTACTGACAGGGAGACACACTCAGATCTGCTGGAGGTGAG 540
DB 481 CCTCAACCTCGAGATAGTACTGACAGGGAGACACACTCAGATCTGCTGGAGGTGAG 540
OY 541 GGACCACTGGAGAGAGAGAAAGCATATGGCGGGATTTCTGAGGGCCAGGATGCTTC 600
DB 541 GGACCACTGGAGAGAGAGAAAGCATATGGCGGGATTTCTGAGGGCCAGGATGCTTC 600
OY 601 CCCAGCGCTGATGGCAGGTGCTTCAGAAAGGTGACTGACTTCAACAAAGGCACTTACCT 660
DB 601 CCCAGCGCTGATGGCAGGTGCTTCAGAAAGGTGACTGACTTCAACAAAGGCACTTACCT 660
OY 661 GGTCACTTCACTGCTGTTGGAGAGGCGCAGGTCTCTCTCTCTGCTGCTCATCCACC 720
DB 661 GGTCACTTCACTGCTGTTGGAGAGGCGCAGGTCTCTCTCTCTGCTGCTCATCCACC 720
OY 721 CAGTGAAGGGGTGTACAGCTCTCTGAGTGAAGCAAGGCTATGACAGGCTGATCT 780
DB 721 CAGTGAAGGGGTGTACAGCTCTCTGAGTGAAGCAAGGCTATGACAGGCTGATCT 780
OY 781 CACTGGCAGTTTGTCAATGGCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 840
DB 781 CACTGGCAGTTTGTCAATGGCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 840
OY 841 CACAAATGCTCAATTTGTCAGTACCTGGACACAGACAGCAAGAGGCTTCTACTGT 900
DB 841 CACAAATGCTCAATTTGTCAGTACCTGGACACAGACAGCAAGAGGCTTCTACTGT 900
OY 901 GAGCCCTCAACACATGGCCCTGTGTCACATGTCACATGTCATGTCATGTCATGTCAT 960
DB 901 GAGCCCTCAACACATGGCCCTGTGTCACATGTCACATGTCATGTCATGTCATGTCAT 960
OY 961 TTCTTATCTTAGCAAAACAAAGAGGCTCTTTGAAAGGTCMAATGCGGTAGAGAT 1020
DB 961 TTCTTATCTTAGCAAAACAAAGAGGCTCTTTGAAAGGTCMAATGCGGTAGAGAT 1020
OY 1021 TATGAAAAATTCATATCAATAGTGTCTCCAAATGCAACACACTGAAAGTCACTGACT 1080
DB 1021 TATGAAAAATTCATATCAATAGTGTCTCCAAATGCAACACACTGAAAGTCACTGACT 1080
OY 1081 GCATGAATCTGGAAAAATTCACACACAGCTTGTGAGATTGATAGAACATCAACAT 1140
DB 1081 GCATGAATCTGGAAAAATTCACACACAGCTTGTGAGATTGATAGAACATCAACAT 1140
OY 1141 CGAGTGGCAAAATATTTGTTATCCCTTGATAGAGATCAATGACATTCACATCAAGAT 1200
DB 1141 CGAGTGGCAAAATATTTGTTATCCCTTGATAGAGATCAATGACATTCACATCAAGAT 1200
OY 1201 GGAGTACCTGACCGCGGCATGACACAACTGAGAGAGAAAAAATCTGTCTATTGTTAT 1260
DB 1201 GGAGTACCTGACCGCGGCATGACACAACTGAGAGAGAAAAAATCTGTCTATTGTTAT 1260
OY 1261 TTCCCTGGGCGACAGATTTCCAGACCTTTCCCATATGATTTTATTCGAAAGGCGCTCAA 1320
DB 1261 TTCCCTGGGCGACAGATTTCCAGACCTTTCCCATATGATTTTATTCGAAAGGCGCTCAA 1320

```

```

OY 1321 TGTCCAAAGGCATTCAGCATCTTCTTGTGAAAGCCGACACATATGTTATCTCAA 1380
DB 1321 TGTCCAAAGGCATTCAGCATCTTCTTGTGAAAGCCGACACATATGTTATCTCAA 1380
OY 1381 AACAGAAACATCCAGGAGATGTACAAATGATCCAGAAAGATTAGTACTTTCATGCTTA 1440
DB 1381 AACAGAAACATCCAGGAGATGTACAAATGATCCAGAAAGATTAGTACTTTCATGCTTA 1440
OY 1441 CATTCATATGATCATATCAATTAAGGACATTTTCCAGATCTGATGATCATTAATGCT 1500
DB 1441 CATTCATATGATCATATCAATTAAGGACATTTTCCAGATCTGATGATCATTAATGCT 1500
OY 1501 CTGGATATATCAATTCATATGAGCAAAATATGTAACCCCTCAACATGTAATGCT 1560
DB 1501 CTGGATATATCAATTCATATGAGCAAAATATGTAACCCCTCAACATGTAATGCT 1560
OY 1561 AATCAGATTAATATATTAATTAATTAATTTGTTAAATTAACAAAAA 1616
DB 1561 AATCAGATTAATATATTAATTAATTAATTTGTTAAATTAACAAAAA 1616

```

## RESULT 2

```

US-09-729-454-30
: Sequence 30. Application US/09729454
: Patient No. US20020137038A1
: GENERAL INFORMATION:
: APPLICANT: Lasek, Amy W.
: APPLICANT: Krasnow, Randi E.
: APPLICANT: Baughn, Mariah R.
: TITLE OF INVENTION: INTESTINAL PROTEINS
: FILE REFERENCE: PC-0028 CIP
: CURRENT APPLICATION NUMBER: US/09/729,454
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PERL Program
: SEQ ID NO 30
: LENGTH: 1005
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. US20020137038A1 GNN 99965027_000007_006
US-09-729-454-30

```

Query Match 62.2%; Score 1005; DH 10; Length 1005;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 593 ATGCTTCCCGAGCGTGATGGAGAGGTGCTTCAAGAAAGTGACTTCAACAAAGGC 652
DB 1 ATGCTTCCCGAGCGTGATGGAGAGGTGCTTCAAGAAAGTGACTTCAACAAAGGC 652
OY 653 ACTTACCTGTCAGCTTCACTGCTGTTCTGGAGAGGCGCAGGTCTCTGCTGCTGCTG 712
DB 61 ACTTACCTGTCAGCTTCACTGCTGTTCTGGAGAGGCGCAGGTCTCTGCTGCTGCTG 120
OY 713 ATCCACCCGAGTGAAGGGGTGTCAGCTCTCTGAGAGTGCAGAAACCAAGGCTATGCAAG 772
DB 121 ATCCACCCGAGTGAAGGGGTGTCAGCTCTCTGAGAGTGCAGAAACCAAGGCTATGCAAG 180
OY 773 GTGATCTTCACTGCGCAGTTTGTCAATGAGCACTTCCAAAGTGCACCTGATGATGAGCTG 832
DB 181 GTGATCTTCACTGCGCAGTTTGTCAATGAGCACTTCCAAAGTGCACCTGATGATGAGCTG 240
OY 833 ATCTTAACCAAAATGCTGAATTTGTCAGTACCTGTCAGCAACAGAGCCAAAGGCTTC 892
DB 241 ATCTTAACCAAAATGCTGAATTTGTCAGTACCTGTCAGCAACAGAGCCAAAGGCTTC 300
OY 893 TACTGTGTAGGCTTCAACACATGCGCTGTGTCACATCACTACATATGATTTCTAAGAC 952
DB 301 TACTGTGTAGGCTTCAACACATGCGCTGTGTCACATCACTACATATGATTTCTAAGAC 360
OY 953 AAGAAAGTTCTTATCTAGCAAAACAAAGAGACCTCTTTGAAAGGTCAATGTGAGT 1012

```

```

Db 361 AAGAAAGTTCTTATCTTAGCAAAAGAAAAGACCTCTTGAAGGTCGAATGTGGGT 420
QY 1013 GTAGAGATTATGCAAAAAATTCATACATTTAGTGTCTGCAAAATGCAACACTGAGTCA 1072
Db 421 GTAGAGATTATGCAAAAAATTCATACATTTAGTGTCTGCAAAATGCAACACTGAGTCA 480
QY 1073 GTGATCTGATGATCTGCAAAATGCAACACAGCTTGTGTGATTTGGATAGAAC 1132
Db 481 GTGATCTGATGATCTGCAAAATGCAACACAGCTTGTGTGATTTGGATAGAAC 540
QY 1133 ATCAACATCCAGTGGCAAAAAATATGTATCCCTGATAGATCATATGACCTATTCAGTC 1192
Db 541 ATCAACATCCAGTGGCAAAAAATATGTATCCCTGATAGATCATATGACCTATTCAGTC 600
QY 1193 AAAGAGATGAGTACCTACCCGGGCACTTGACAGAACTGGAGAGAAAAAAATACTGTC 1252
Db 601 AAAGAGATGAGTACCTACCCGGGCACTTGACAGAACTGGAGAGAAAAAAATACTGTC 660
QY 1253 ATGTATTTGCTGGGCGCAGCATTTCCAGCCCTTCCCATGATGTTTATTCGGAAG 1312
Db 661 ATGTATTTGCTGGGCGCAGCATTTCCAGCCCTTCCCATGATGTTTATTCGGAAG 720
QY 1313 GCGCTAATGTCCAAAGCCATTCAGCATCTCTTGTGAAAGCCGACAGCATATGTT 1372
Db 721 GCGCTAATGTCCAAAGCCATTCAGCATCTCTTGTGAAAGCCGACAGCATATGTT 780
QY 1373 ATCATCAAAACAGAAACATCAGGAGATGACAAATGATGCAAAAGATTTAGTACTTT 1432
Db 781 ATCATCAAAACAGAAACATCAGGAGATGACAAATGATGCAAAAGATTTAGTACTTT 840
QY 1433 CATGTTACATTCATATTCATCATTAAGACATTTTCAGAGATCCAGTGTAGTATC 1492
Db 841 CATGTTACATTCATATTCATCATTAAGACATTTTCAGAGATCCAGTGTAGTATC 900
QY 1493 ATTGATGCGTGGATATACAAATGTCATATGCGCAATATATACCCACTCAACAT 1552
Db 901 ATTGATGCGTGGATATACAAATGTCATATGCGCAATATATACCCACTCAACAT 960
QY 1553 GTAGTGGAAATGAGTATATATTTAATTAACATATTTGTTAA 1597
Db 961 GTAGTGGAAATGAGTATATATTTAATTAACATATTTGTTAA 1005

RESULT 3
US-09-981-353-144
Sequence 144, Application US/09981353
Patent No. US20020160382A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy W.
TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
FILE REFERENCE: PA-0038 US
CURRENT APPLICATION NUMBER: US/09/981,353
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 144
LENGTH: 1212
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020160382A1 1397029.1
US-09-981-353-144

Query Match 34.7%, Score 560, DB 9, Length 1212;
Best Local Similarity 99.7%, Pred. No. 2, 2e-232;
Matches 730; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 61 GAGGTATGCCAGTAGACACTGTTGGCTTCAGAAAGAAATTCACACCTAGCTGCCAG 120
Db 61 GAGGTATGCCAGTAGACACTGTTGGCTTCAGAAAGAAATTCACACCTAGCTGCCAG 120
QY 121 AGAGTCTATGTATGGGATTTGAAATCTGTAACTTAAAGGATCTCAATCAAGAAATAG 180
Db 121 AGAGTCTATGTATGGGATTTGAAATCTGTAACTTAAAGGATCTCAATCAAGAAATAG 180
QY 181 TATGATTAATTTAAGTCACT -ATTGGACGTGTTGTTATATTAGCTCTGTGATCATTT 239
Db 181 TATGATTAATTTAAGTCACT -ATTGGACGTGTTGTTATATTAGCTCTGTGATCATTT 240
QY 240 TTACAGTTTCCAGAACTCCACAAGGTTGGTCTGTCTTAACTTATTCATCTCCCTCC 299
Db 241 TTACAGTTTCCAGAACTCCACAAGGTTGGTCTGTCTTAACTTATTCATCTCCCTCC 300
QY 300 ATTACTGGAACAACTCCAAAAGTCTTATTCCTTAAACACCACTGATATCATTTAAG 359
Db 301 ATTACTGGAACAACTCCAAAAGTCTTATTCCTTAAACACCACTGATATCATTTAAG 360
QY 360 CACTAACAGAGACTGAAGCTGAGATTTAAAGGAAATCATAGAGAACTAGATGACATCC 419
Db 361 CACTAACAGAGACTGAAGCTGAGATTTAAAGGAAATCATAGAGAACTAGATGACATCC 420
QY 420 CACCCAGACTTTTACCCAGGTGAACACCCACACGACGACACATAGACACCA 479
Db 421 CACCCAGACTTTTACCCAGGTGAACACCCACACGACGACACATAGACACCA 480
QY 480 TCCTCAACCTCGAGATAGTACTAGTGAAGGAGACAGCTGCAATCTGTGTAGAGTGA 539
Db 481 TCCTCAACCTCGAGATAGTACTAGTGAAGGAGACAGCTGCAATCTGTGTAGAGTGA 540
QY 540 GGGACACTGTGGAGCCGAGAAACAAATATGCGGGGATTTCTGAGGGCCAGATGCTT 599
Db 541 GGGACACTGTGGAGCCGAGAAACAAATATGCGGGGATTTCTGAGGGCCAGATGCTT 600
QY 600 CCCAGCCGCTGATGGCAGTGTCTTCAGGAAAGTGACTGACTTCAACACAGCCACTACC 659
Db 601 CCCAGCCGCTGATGGCAGTGTCTTCAGGAAAGTGACTGACTTCAACACAGCCACTACC 660
QY 660 TGGTCAAGCTTACCTGTCTTCTGGAGGCGCAGGTCTCTGTCTGTGTGCTATCCACC 719
Db 661 TGGTCAAGCTTACCTGTCTTCTGGAGGCGCAGGTCTCTGTCTGTGTGCTATCCACC 720
QY 720 CCAAGTGAAGGG 731
Db 721 CCAAGTGAAGGG 732

RESULT 4
US-09-729-454-5/c
Sequence 5, Application US/09729454
Patent No. US20020137038A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy W.
TITLE OF INVENTION: INTESTINAL PROTEINS
FILE REFERENCE: PC-0028 CIP
CURRENT APPLICATION NUMBER: US/09/729,454
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 5
LENGTH: 621
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020137038A1 71851705V1
US-09-729-454-5

```



Query Match	34.5%	Score 556;	DB 10;	Length 621;
Best Local Similarity	99.8%	Pred. No. 1.7e+231;		
Matches 608; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	210	TGCTTTTATATAGGCTCCGCGATGCAATTTTTCAGATTTTCACAACTCCACAAAGGTTT	26
Db	621	TGTTTATTATATAGGCTCCGCGATGCAATTTTTCAGATTTTTCACAACTCCACAAAGGTTT	56
QY	270	GGCTGCTCTTAACCTTATATCCATCTCCCTCCATTACTGGAACACTCCACAAAGTCTTAT	32
Db	561	GGTCTGCTCTTAACCTTATATCCATCTCCCTCCATTACTGGAACACTCCACAAAGTCTTAT	50
QY	330	TCCCTTAAACACCACTGATATCATTTAAAGCCATTAAGCCATTAACGACATGAACTCAGATTAAGG	38
Db	501	TCCCTTAAACACCACTGATATCATTTAAAGCCATTAACGACATGAACTCAGATTAAGG	44
QY	390	AAATCATAGAGAAACTAGATGATGAGATGCCACCCAGACCTTTTCACCCAGTGAACACCA	44
Db	441	AAATCATAGAGAAACTAGATGATGAGATGCCACCCAGACCTTTTCACCCAGTGAACACCA	38
QY	450	CCACCAGGGCCACACATAGCACAGCCACACTCCCTAACCCCTCGAGATAGCTACTGACAGG	50
Db	381	CCACCAGGGCCACACATAGCACAGCCACACTCCCTAACCCCTCGAGATAGCTACTGACAGG	32
QY	510	GAGACACAGCTGCACATCTCTGTGAGAGGTGAGGGACCACTTGGAGCGCAGGAGACAAATATG	56
Db	321	GAGACACAGCTGCACATCTCTGTGAGAGGTGAGGGACCACTTGGAGCGCAGGAGACAAATATG	26
QY	570	GGGGGGAATTTCTGAGGGGCGACAGATGTCTTCCCGCAGCCGTATGCGAGGTGCTTAGGAA	62
Db	261	GGGGGGAATTTCTGAGGGGCGACAGATGTCTTCCCGCAGCCGTATGCGAGGTGCTTAGGAA	20
QY	630	AGGTGACTGACTTCAACAAGGCGACCTTACTGGTCAAGCTTCTGTGTTCTGSGAGGGCC	68
Db	201	AGGTGACTGACTTCAACAAGGCGACCTTACTGGTCAAGCTTCTGTGTTCTGSGAGGGCC	14
QY	690	AGGTCTCTGTGCTCTGCTGCTCATCATCCACCCAGTGAAGGGGTGTGAGCTCTGTGAGTG	74
Db	141	AGGTCTCTGTGCTCTGCTGCTCATCATCCACCCAGTGAAGGGGTGTGAGCTCTGTGAGTG	82
QY	750	CAGGACACCAAGGCTATGACAGAGGTGATCTTCACTAGTGGCCACTTTGTCAATGCGACTTCC	80
Db	81	CAGGACACCAAGGCTATGACAGAGGTGATCTTCACTAGTGGCCACTTTGTCAATGCGACTTCC	22
QY	810	AAGTCCACT 818	
Db	21	AAGTCCACT 13	
RESULT 5			
US-09-981-353-145			
Sequence 145, Application US/09981353			
Patent No. US20020160382A1			
GENERAL INFORMATION:			
APPLICANT: Laasek, Amy W.			
APPLICANT: Jones, David A.			
TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER			
FILE REFERENCE: PA-0038 US			
CURRENT APPLICATION NUMBER: US/09/981,353			
NUMBER OF SEQ ID NOS: 194			
SOFTWARE: PERL Program			
SEQ ID NO 145			
LENGTH: 841			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: misc_feature			
OTHER INFORMATION: Incyte ID No. US20020160382A1 403560.1			
US-09-981-353-145			

Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

[illegible]

```

RESULT 6
US-09-729-454-6/C
: Sequence 6, Application US/09729454
: Patient No. US020020137038A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy W.
APPLICANT: Krasnow, Randi E.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: INTESTINAL PROTEINS
FILE REFERENCE: PC-0028 CIP
CURRENT APPLICATION NUMBER: US/09/729,454
CURRENT FILING DATE: 2000-12-04
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL program
SEQ ID NO 6
LENGTH: 545
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: incyle ID No. US20020137038A1 70255757V1
US-09-729-454-6

```

Query Match	25 38	Score 409	DB 107	Length 545
Best Local Similarity	99.68	Pred No 3.30e-167		
Matches 509	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY 319	AAATGCTTATTCCTTAAACACACACTGATATCTTTAAAGCCATACAGAGACTGAAGT	378		
DB 511	AAAGTCTTATTCCTTAAACACACACTGATATCTTTAAAGCCATACAGAGACTGAAGT	452		
QY 379	CAGATTAAGGAATCATATAGAGAACTGATGACACGATCCACCCAGACTTTATACCA	448		
DB 451	CAGATTAAGGAATCATATAGAGAACTGATGACACGATCCACCCAGACTTTATACCA	444		

```
OY 439 CGTGAACACACACACACACACACATAGACAGCCACCATCTCAACCCCTGAGATAC 498
DB 391 CGTGAACACACACACACACACACACACACACACACACACACACACACACACAC 332
OY 499 GTACTGACAGGAGACACACACACACACACACACACACACACACACACACACAC 558
DB 331 GTACTGACAGGAGACACACACACACACACACACACACACACACACACACACAC 272
OY 559 GAAATATATGAGGAGATTCCTGAGGAGACACACACACACACACACACACACAC 618
DB 271 GAAATATATGAGGAGATTCCTGAGGAGACACACACACACACACACACACACAC 212
OY 619 TGCTTCAGCAAAAGTACTGACTTCAACACACACACACACACACACACACACAC 678
DB 211 TGCTTCAGCAAAAGTACTGACTTCAACACACACACACACACACACACACACAC 152
OY 679 CTGGACAGCCACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 738
DB 151 CTGGACAGCCACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 92
OY 739 TCTCTGAGTGTGACAGCAACACACACACACACACACACACACACACACACAC 798
DB 91 TCTCTGAGTGTGACAGCAACACACACACACACACACACACACACACACACAC 32
OY 799 TGGCACTTCCCAAGTCCAGCTCTGATGTGGC 829
DB 31 TGGCACTTCCCAAGTCCAGCTCTGATGTGGC 1
```

## RESULT 7

```
US-09-729-454-9
; Sequence 9, Application US/09729454
; Patent No. US20020137038A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Haughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 7128544H1
US-09-729-454-9
```

```
Query Match 23.2% Score 375; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 1.5e-152;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1221 TTGACGAACTGAGGAGAAAAAATATGTCATGTATTTCCCTGGCCACGATTTC 1280
DB 20 TTGACGAACTGAGGAGAAAAAATATGTCATGTATTTCCCTGGCCACGATTTC 79
OY 1281 GACCCCTTCCATGATGATTTTATCCGAGGGCCCTCAATGTCACAAAGCAATTGAG 1340
DB 80 GACCCCTTCCATGATGATTTTATCCGAGGGCCCTCAATGTCACAAAGCAATTGAG 139
OY 1341 ATCTTCTTGTAGAACCCAGACATGCTTATCATCAAAAACAAACATCGAGAGA 1400
DB 140 ATCTTCTTGTAGAACCCAGACATGCTTATCATCAAAAACAAACATCGAGAGA 199
OY 1401 TGTACATGATGACGAAAGATTTAGTACTTTCATGTTACATTCATATATCTATTA 1460
DB 200 TGTACATGATGACGAAAGATTTAGTACTTTCATGTTACATTCATATATCTATTA 259
```

```
OY 1461 AGGACATTTTCAGAGATCTCAGTGTAGATATCATGATGCGGTGGATATTAACATTGAT 1520
DB 260 AGGACATTTTCAGAGATCTCAGTGTAGATATCATGATGCGGTGGATATTAACATTGAT 319
OY 1521 ATGGACAAATATATGACACCCACCTCAACATGTAGTGGAAATACATATATATAT 1580
DB 320 ATGGACAAATATATGACACCCACCTCAACATGTAGTGGAAATACATATATATATAT 379
OY 1581 TAAACTATATTTGTT 1595
DB 380 TAAACTATATTTGTT 394
```

## RESULT 8

```
US-09-729-454-4
; Sequence 4, Application US/09729454
; Patent No. US20020137038A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Haughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 2771041H1
US-09-729-454-4
```

```
Query Match 14.9% Score 240; DB 10; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.8e-94;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ATCCATGCTAAAGTAAACAACTGCAACTATATATGCAATTTTGTATAGACAA 60
DB 1 ATCCATGCTAAAGTAAACAACTGCAACTATATATGCAATTTTGTATAGACAA 60
OY 61 GAGGTATGCGAGTACAGACAGTGGGCTTACAGAGAAATTCACACCTAGCTGCCAG 120
DB 61 GAGGTATGCGAGTACAGACAGTGGGCTTACAGAGAAATTCACACCTAGCTGCCAG 120
OY 121 AGACTCTATGTATGAGATTGACAACTCTAAACTAAAGATCCTAATCATGAAATAG 180
DB 121 AGACTCTATGTATGAGATTGACAACTCTAAACTAAAGATCCTAATCATGAAATAG 180
OY 181 TATGATTAATTAATGATACATATGACATGTTGTTATATATAGCCTCTGATCATTT 240
DB 181 TATGATTAATTAATGATACATATGACATGTTGTTATATATAGCCTCTGATCATTT 240
```

## RESULT 9

```
US-09-729-454-7
; Sequence 7, Application US/09729454
; Patent No. US20020137038A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Haughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 236
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 5596934H1
; NAME/KEY: unsure
; LOCATION: 228
; OTHER INFORMATION: a, t, c, g, or other
US-09-729-454-7

```

```

Query Match      13.2%; Score 213; DB 10; Length 236;
Best Local Similarity 100.0%; Pred. No. 1.2e-82;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 766 TGACAGGTGATCTTCACTGGCCAGTTGTGCATGCGACCTTCCCAAGTCCACTGTGAATG 825
    |||||||
DB 15 TGACAGGTGATCTTCACTGGCCAGTTGTGCATGCGACCTTCCCAAGTCCACTGTGAATG 74
    |||||||
QY 826 TGGCTTGATCTTAACACAAATGCTGAATGTGCCAGTACCTGGACACAGACCAAGA 885
    |||||||
DB 75 TGGCTTGATCTTAACACAAATGCTGAATGTGCCAGTACCTGGACACAGACCAAGA 134
    |||||||
QY 886 AGGCTTCTACTGTGTGAGGCTCAACACATGCGCTGTCTGCTCAGTCTACATGTATTC 945
    |||||||
DB 135 AGGCTTCTACTGTGTGAGGCTCAACACATGCGCTGTCTGCTCAGTCTACATGTATTC 194
    |||||||
QY 946 TAAGAACAGAAAGTTCTTATCTTAGCAACA 978
    |||||||
DB 195 TAAGAACAGAAAGTTCTTATCTTAGCAACA 227
    |||||||

```

```

RESULT 10
US-10-046-935-181
; Sequence 181, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aljun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-181

```

```

US-10-046-935-181

```

```

Query Match      12.3%; Score 198; DB 9; Length 554;
Best Local Similarity 100.0%; Pred. No. 3.6e-76;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 864 ACCTGACACAGACAGACAGAGGCTTCTACTGTGTGAGGCTCAACACATGCCCTGTG 923
    |||||||
DB 1 ACCTGACACAGACAGACAGAGGCTTCTACTGTGTGAGGCTCAACACATGCCCTGTG 60
    |||||||
QY 924 CTGACATGCTACATGATGATTTCTAAGAACAGAAAGTTCTTATCTTAGCAACAAGAA 983
    |||||||
DB 61 CTGACATGCTACATGATGATTTCTAAGAACAGAAAGTTCTTATCTTAGCAACAAGAA 120
    |||||||
QY 984 AGAGCTCTTTGAAGAGTCAATGTGGGTGATGAGATTTATGAAAAATTCATACAAATA 1043
    |||||||
DB 121 AGAGCTCTTTGAAGAGTCAATGTGGGTGATGAGATTTATGAAAAATTCATACAAATA 180
    |||||||
QY 1044 GTGTCTCAAAATGCAACA 1061
    |||||||
DB 181 GTGTCTCAAAATGCAACA 198
    |||||||

```

```

RESULT 11
US-10-046-935-388/C
; Sequence 388, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aljun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 388
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 546-560
; OTHER INFORMATION: n = A,T,C or G
US-10-046-935-388

```

```

Query Match      5.8%; Score 93; DB 9; Length 580;
Best Local Similarity 99.3%; Pred. No. 7.3e-31;
Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1062 CACTGAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1121
    |||||||
DB 144 CACTGAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 85
    |||||||
QY 1122 TGGATGAGACATCAATCACTGATGATGATGATGATGATGATGATGATGATGATG 25
    |||||||
DB 84 TGGATGAGACATCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATG 25
    |||||||
QY 1182 CCTATTGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1205
    |||||||
DB 24 CCTATTGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1
    |||||||

```

```

RESULT 12
US-09-729-454-11
; Sequence 11, Application US/09729454
; Patent No. US20020137038A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Krasnow, Randi E.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: INTESTINAL PROTEINS
; FILE REFERENCE: PC-0028 CIP
; CURRENT APPLICATION NUMBER: US/09/729,454
; CURRENT FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 661
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137038A1 1339495F6
US-09-729-454-11

```

```

Query Match      5.4%; Score 88; DB 10; Length 661;
Best Local Similarity 100.0%; Pred. No. 1e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 351 CATTAAAGCCACTAACAGAGACTGACTGAGATTAAGGAATCATAGAGAACTAGATC 410  
|||||  
Db 344 CATTAAAGCCACTAACAGAGACTGACTGAGATTAAGGAATCATAGAGAACTAGATC 403  
|||||  
QY 411 AGCAGATCCAGCCAGACCTTTCACCCA 438  
|||||  
Db 404 AGCAGATCCAGCCAGACCTTTCACCCA 431

## RESULT 13

US-09-729-454-31  
Sequence 31, Application US/09729454  
Patent No. US20020137038A1  
GENERAL INFORMATION:

APPLICANT: Iasek, Amy W.  
APPLICANT: Krasnow, Randi E.  
APPLICANT: Haughn, Mariah R.  
TITLE OF INVENTION: INTESTINAL PROTEINS  
FILE REFERENCE: PC-0028 CIP  
CURRENT APPLICATION NUMBER: US/09/729,454  
CURRENT FILING DATE: 2000-12-04  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PERL program  
SEQ ID NO 31

LENGTH: 1545

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US20020137038A1 GNN.g9795680\_006.edlt  
US-09-729-454-31

## Query Match

Host Local Similarity 5.4%; Score 88; DB 10; Length 1545;  
Best Local Similarity 100.0%; Pred. No. 1e-28;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 CATTAAAGCCACTAACAGAGACTGACTGAGATTAAGGAATCATAGAGAACTAGATC 410  
|||||  
Db 86 CATTAAAGCCACTAACAGAGACTGACTGAGATTAAGGAATCATAGAGAACTAGATC 145  
|||||  
QY 411 AGCAGATCCAGCCAGACCTTTCACCCA 438  
|||||  
Db 146 AGCAGATCCAGCCAGACCTTTCACCCA 173

## RESULT 14

US-09-729-454-10  
Sequence 10, Application US/09729454  
Patent No. US20020137038A1  
GENERAL INFORMATION:

APPLICANT: Iasek, Amy W.  
APPLICANT: Krasnow, Randi E.  
APPLICANT: Haughn, Mariah R.  
TITLE OF INVENTION: INTESTINAL PROTEINS  
FILE REFERENCE: PC-0028 CIP  
CURRENT APPLICATION NUMBER: US/09/729,454  
CURRENT FILING DATE: 2000-12-04  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PERL program  
SEQ ID NO 10

LENGTH: 2248

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US20020137038A1 7484349CB1  
US-09-729-454-10

## Query Match

Host Local Similarity 5.4%; Score 88; DB 10; Length 2248;  
Best Local Similarity 100.0%; Pred. No. 1e-28;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 344 CATTAAAGCCACTAACAGAGACTGACTGAGATTAAGGAATCATAGAGAACTAGATC 403  
|||||  
QY 411 AGCAGATCCAGCCAGACCTTTCACCCA 438  
|||||  
Db 404 AGCAGATCCAGCCAGACCTTTCACCCA 431

## RESULT 15

US-09-729-454-8  
Sequence 8, Application US/09729454  
Patent No. US20020137038A1  
GENERAL INFORMATION:

APPLICANT: Iasek, Amy W.  
APPLICANT: Krasnow, Randi E.  
APPLICANT: Haughn, Mariah R.  
TITLE OF INVENTION: INTESTINAL PROTEINS  
FILE REFERENCE: PC-0028 CIP  
CURRENT APPLICATION NUMBER: US/09/729,454  
CURRENT FILING DATE: 2000-12-04  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PERL program  
SEQ ID NO 8

LENGTH: 414

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

OTHER INFORMATION: Incyte ID No. US20020137038A1 3229449P6  
LOCATION: 47, 105, 248

OTHER INFORMATION: a, t, c, g, or other  
US-09-729-454-8

## Query Match

Host Local Similarity 4.5%; Score 73; DB 10; Length 414;  
Best Local Similarity 100.0%; Pred. No. 3.1e-22;

Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1057 CAACAGACTGAAGTCAGTGCATGATCTGCAAAATTGCAACACAGCTTCTGT 1116  
|||||  
Db 175 CAACAGACTGAAGTCAGTGCATGATCTGCAAAATTGCAACACAGCTTCTGT 234  
|||||  
QY 1117 GGATTTGGATAGG 1129  
|||||  
Db 235 GGATTTGGATAGG 247

Search completed: November 22, 2002, 04:45:20  
Job time : 873 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 01:11:26 ; Search time 4129 Seconds

(without alignments)  
11390.200 Million cell updates/sec

Title: US-09-729-454-3

Perfect score: 1616  
Sequence: 1 atccatgcctaaaggaataca.....aatacaaaaaaaaaaaaaa 1616

Scoring table:

Gapop 60.0 , Capext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

GenEmbl:.\*  
1: gb\_poa.\*  
2: gb\_hlg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_hlg\_hum.\*  
31: em\_hlg\_liv.\*  
32: em\_hlg\_other.\*  
33: em\_hlg\_mus.\*  
34: em\_hlg\_pin.\*  
35: em\_hlg\_rtd.\*  
36: em\_hlg\_mam.\*  
37: em\_hlg\_vrl.\*  
38: em\_sy.\*  
39: em\_higo\_hum.\*  
40: em\_higo\_mus.\*  
41: em\_higo\_other.\*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1616	100.0	1616	6	AX458378 Sequence
2	1005	62.2	1005	6	AX458405 Sequence
3	987	61.1	1739	6	AX358758 Sequence
4	987	61.1	1739	6	AX362251 Sequence
5	983	60.8	1138	6	AX383510 Sequence
6	738	45.7	81975	2	AP000450 Sequence
7	738	45.7	146841	2	AC020549 Homo sapi
8	738	45.7	187852	2	AC068041 Homo sapi
9	558	34.5	621	6	AX458380 Sequence
10	528	32.7	1704	9	AX000134 Sequence
11	409	26.3	545	6	AX458381 Sequence
12	400	24.8	400	6	AX386470 Sequence
13	375	23.2	394	6	AX458384 Sequence
14	350	21.7	178494	2	AC067812 Homo sapi
15	240	14.9	240	6	AX458379 Sequence
16	213	13.2	236	6	AX458382 Sequence
17	198	12.3	557	6	AX339344 Sequence
18	192	11.9	275	6	AX079818 Sequence
19	181	11.2	532	6	AX351427 Sequence
20	93	5.8	583	6	AX340141 Sequence
21	88	5.4	211	6	AX385391 Sequence
22	88	5.4	611	6	AX385432 Sequence
23	88	5.4	661	6	AX458386 Sequence
24	88	5.4	1545	6	AX458406 Sequence
25	88	5.4	1953	6	BC029049 Homo sapi
26	88	5.4	2248	6	AX458385 Sequence
27	88	5.4	81975	2	AP000450 Sequence
28	83	5.1	2222	9	AK057953 Sequence
29	83	5.1	143590	9	AP002774 Homo sapi
30	73	4.5	414	6	AX458383 Sequence
31	47	2.9	518	6	AX458387 Sequence
32	42	2.6	2144	4	OC000000 Sequence
33	34	2.1	244	6	AX458393 Sequence
34	34	2.1	270	6	AX458394 Sequence
35	34	2.1	291	6	AX458392 Sequence
36	33	2.0	60942	2	AC124611 Sequence
37	32	2.0	206	6	AX458391 Sequence
38	32	2.0	257	6	AX458397 Sequence
39	29	1.8	58145	2	AC125828 Rattus no
40	29	1.8	178494	2	AC067812 Homo sapi
41	25	1.5	288	6	AX458395 Sequence
42	24	1.5	275	6	AX458396 Sequence
43	24	1.5	394	6	AX458384 Sequence
44	24	1.5	150489	2	AC068854 Homo sapi
45	24	1.5	165028	9	AC110758 Homo sapi

## ALIGNMENTS

RESULT 1  
AX458378  
LOCUS AX458378 1616 bp DNA linear PAT 08-JUL-2002  
DEFINITION Sequence 3 from Patent WO0246411.  
ACCESSION AX458378  
VERSION AX458378.1 GI:21725046  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 Lasek,A.W., Krasnow,R. and Baughn,M.R.  
AUTHORS Intestinal proteins  
TITLE Patent: WO 0246411-A 3 13-JUN-2002;  
JOURNAL

Incyle Genomes, Inc. (US)  
 Location/Qualifiers  
 1..1616  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="Incyle ID No: 3229449CB1"

BASE COUNT 511 a 374 c 324 g 407 t  
 ORIGIN

Query Match 100.0%; Score 1616; DB 6; Length 1616;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATCCATGCTAAAGCTAAACAACTGCAACTTATCTGCAATTTATTTGGTATAGACA 60
DB 1 ATCCATGCTAAAGCTAAACAACTGCAACTTATCTGCAATTTATTTGGTATAGACA 60
QY 61 GAGGTATGCCAGTAGACACAGTGGTGGCTTCAGAAATTTCTCAACACCTAGCTGCCAG 120
DB 61 GAGGTATGCCAGTAGACACAGTGGTGGCTTCAGAAATTTCTCAACACCTAGCTGCCAG 120
QY 121 AGAGCTATATGATGATGAAACAATCTGTAAACTAAAGATCCTAATCATGAAATAG 180
DB 121 AGAGCTATATGATGATGAAACAATCTGTAAACTAAAGATCCTAATCATGAAATAG 180
QY 181 TATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 TATGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 TACAGTTTCCAGACCTCCAGCAAGGTTGGTGTCTCTAACTTATCATCTCCCTCCA 300
DB 241 TACAGTTTCCAGACCTCCAGCAAGGTTGGTGTCTCTAACTTATCATCTCCCTCCA 300
QY 301 TTACTGCAACACTCCAGCAAGGTTGGTGTCTCTAACTTATCATCTCCCTCCA 360
DB 301 TTACTGCAACACTCCAGCAAGGTTGGTGTCTCTAACTTATCATCTCCCTCCA 360
QY 361 ACTAAGAGAGACTGAAGCTGATTAAGAAATCATAGAACTAGATGACAGATCCC 420
DB 361 ACTAAGAGAGACTGAAGCTGATTAAGAAATCATAGAACTAGATGACAGATCCC 420
QY 421 ACCGAGACTTTCCAGCAGTGAACACACACAGCCGACAGATAGACAGCCACAT 480
DB 421 ACCGAGACTTTCCAGCAGTGAACACACACAGCCGACAGATAGACAGCCACAT 480
QY 481 CCTCAACCTCGAGATGCTACTGACAGGAGACACGCTGACATCTCTGTGAGGTAG 540
DB 481 CCTCAACCTCGAGATGCTACTGACAGGAGACACGCTGACATCTCTGTGAGGTAG 540
QY 541 GGACCACTTGGAGAGCAGAAATATGGCGGGATTCTCTGAGGGCCAGATGCTTC 600
DB 541 GGACCACTTGGAGAGCAGAAATATGGCGGGATTCTCTGAGGGCCAGATGCTTC 600
QY 601 CCCAGGGTATGAGAGTGGCTTCAGAAAGGTGACTACTTCAACAAAGGCACTACT 660
DB 601 CCCAGGGTATGAGAGTGGCTTCAGAAAGGTGACTACTTCAACAAAGGCACTACT 660
QY 661 GGTCAGCTTCACTCTTCTGAGGAGGGCCAGGTCTCTGTCTGTCTGTCTATCCACC 720
DB 661 GGTCAGCTTCACTCTTCTGAGGAGGGCCAGGTCTCTGTCTGTCTGTCTATCCACC 720
QY 721 CAGTGAAGGGGTGTGAGCTCTCTGAGTGAAGAAACCAAGCTATAGAGGGTATCTT 780
DB 721 CAGTGAAGGGGTGTGAGCTCTCTGAGTGAAGAAACCAAGCTATAGAGGGTATCTT 780
QY 781 CACTGGCAAGTTTGTCAATGGCACTTCCCAAGTCCACTCTCAATAGTGGCTGATCTTAA 840
DB 781 CACTGGCAAGTTTGTCAATGGCACTTCCCAAGTCCACTCTCAATAGTGGCTGATCTTAA 840
QY 841 CACAATAGCTGAATGTGCGCACTACTGACAGACAGACAGCAAGAGGCTTCACTGTGT 900
DB 841 CACAATAGCTGAATGTGCGCACTACTGACAGACAGACAGCAAGAGGCTTCACTGTGT 900
  
```

```

QY 901 GAGGCTCAACACATGCCCTGTGCTGCACTACCTACATGATTTCTTAAGACAGAAAGT 960
DB 901 GAGGCTCAACACATGCCCTGTGCTGCACTACCTACATGATTTCTTAAGACAGAAAGT 960
QY 961 TTCTTATCTTACCAACACAGAAAGAGCCTCTTTTGAAGGTCAATGTGGGTGATAGAT 1020
DB 961 TTCTTATCTTACCAACACAGAAAGAGCCTCTTTTGAAGGTCAATGTGGGTGATAGAT 1020
QY 1021 TATGAAAAATTCATACATTAATGATGCTGCTCCAAATGCAACAGCACTGAATGAGATCT 1080
DB 1021 TATGAAAAATTCATACATTAATGATGCTGCTCCAAATGCAACAGCACTGAATGAGATCT 1080
QY 1081 GCATGAATCTGAAATTTGCAACACAGCTTCTGTGATTTGATATGAAATATCAATAT 1140
DB 1081 GCATGAATCTGAAATTTGCAACACAGCTTCTGTGATTTGATATGAAATATCAATAT 1140
QY 1141 CCAATGCAAAATATTTGATATGCTGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 CCAATGCAAAATATTTGATATGCTGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 CGAGTACTTACCCGGGCAATTCAGACAACTGACAGCAUAAAAAATCTGTCAATTCAT 1260
DB 1201 CGAGTACTTACCCGGGCAATTCAGACAACTGACAGCAUAAAAAATCTGTCAATTCAT 1260
QY 1261 TTCCCTGGGCGAGATTTCAAGACCTTTCCGATGATTTTATCCGAAGGGCCCTGAA 1320
DB 1261 TTCCCTGGGCGAGATTTCAAGACCTTTCCGATGATTTTATCCGAAGGGCCCTGAA 1320
QY 1321 TGTCCACAAAGCCATTCAGATCTCTCTGAGAAAGCCACAGCACTATGCTATCAATCA 1380
DB 1321 TGTCCACAAAGCCATTCAGATCTCTCTGAGAAAGCCACAGCACTATGCTATCAATCA 1380
QY 1381 AACGAAAAACATCAGGAGATGATCAATATGACAGAAAGATTTAGTCACTTCACTGTA 1440
DB 1381 AACGAAAAACATCAGGAGATGATCAATATGACAGAAAGATTTAGTCACTTCACTGTA 1440
QY 1441 CATTCATATCTCATATCAATTAAGGACATTTCCAGATCTCAGTGTGATGATGATGATG 1500
DB 1441 CATTCATATCTCATATCAATTAAGGACATTTCCAGATCTCAGTGTGATGATGATGATG 1500
QY 1501 CTGGATATACAAATTCGATGTCAGCAAAATATATGACCCCACTCAAGATGATGATG 1560
DB 1501 CTGGATATACAAATTCGATGTCAGCAAAATATATGACCCCACTCAAGATGATGATG 1560
QY 1561 AAATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1616
DB 1561 AAATCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1616
  
```

RESULT 2  
 AX458405 1005 bp DNA linear PAT 08 JUN 2002  
 LOCUS AX458405  
 DEFINITION Sequence 30 from Patent WO0246411.  
 ACCESSION AX458405  
 VERSION AX458405.1 GI:21725073  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 Lasak, A.W., Krasnow, R. and Baughm, M.R.  
 TITLE Intestinal proteins  
 JOURNAL Patent: WO 0246411-A 30 13 JUN 2002;  
 Incyle Genomes, Inc. (US)  
 Location/Qualifiers  
 source  
 1..1005  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="Incyle ID No: GNN.g9965027\_000007\_006"

BASE COUNT 303 a 225 c 211 g 266 t  
 ORIGIN

Query Match 62.2%; Score 1005; DB 6; Length 1005;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 593 ATGCTGCCCAAGCGCTGATGCGAGTGTCTGAGAAAAGTACTGACTTCAACAGCGC 652  
DB 1 ATGCTGCCCAAGCGCTGATGCGAGTGTCTGAGAAAAGTACTGACTTCAACAGCGC 60

QY 653 ACCTACTGCTGACCTTCACTGCTTGTGGAGGGCCAGCTGCTGCTGCTGCTGCTG 712  
DB 61 ACCTACTGCTGACCTTCACTGCTTGTGGAGGGCCAGCTGCTGCTGCTGCTGCTGCTG 120

QY 713 ATCCAGCCCAAGTGAAGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 772  
DB 121 ATCCAGCCCAAGTGAAGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY 773 GTGATCTTCACTGCGCGAGTTTGTCAATGCGACTTCCCAATGCGACTGAAATGCGCTG 832  
DB 181 GTGATCTTCACTGCGCGAGTTTGTCAATGCGACTTCCCAATGCGACTGAAATGCGCTG 240

QY 833 ATCTTAACACAAATGCTGAATTTGTGCGAGTCTGACACAGACAGACCAAGAGCTTC 892  
DB 241 ATCTTAACACAAATGCTGAATTTGTGCGAGTCTGACACAGACAGACCAAGAGCTTC 300

QY 893 TACTGTGAGGCTGCAACATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952  
DB 301 TACTGTGAGGCTGCAACATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 953 AAGAAAGTTTCTTATCTTGAACAACAAGAGGCTCTTTGAAAGCTCAATGTGGCT 1012  
DB 361 AAGAAAGTTTCTTATCTTGAACAACAAGAGGCTCTTTGAAAGCTCAATGTGGCT 420

QY 1013 GTAGAGATTATGAAAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTA 1072  
DB 421 GTAGAGATTATGAAAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTA 480

QY 1073 GTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132  
DB 481 GTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 1133 ATCAACATCCAGTGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1192  
DB 541 ATCAACATCCAGTGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600

QY 1193 AAGAGATGAGTACTGACCGCGGCGCATTTGACAGAACTGAGAGAGAGAGAGAGAGAG 1252  
DB 601 AAGAGATGAGTACTGACCGCGGCGCATTTGACAGAACTGAGAGAGAGAGAGAGAGAG 660

QY 1253 ATGTTATTTTCTGCGCGCATTTGACAGAACTGAGAGAGAGAGAGAGAGAGAGAG 1312  
DB 661 ATGTTATTTTCTGCGCGCATTTGACAGAACTGAGAGAGAGAGAGAGAGAGAGAG 720

QY 1313 GCGCTCAATGTCACAAAGCCATTGACAGATCTTCTGAGAGAGAGAGAGAGAGAGAG 1372  
DB 721 GCGCTCAATGTCACAAAGCCATTGACAGATCTTCTGAGAGAGAGAGAGAGAGAGAG 780

QY 1373 ATCATCAAAAAGCAAAATCATCGAGAGATGATGATGATGATGATGATGATGATGATG 1432  
DB 781 ATCATCAAAAAGCAAAATCATCGAGAGATGATGATGATGATGATGATGATGATGATG 840

QY 1433 CATGTTTACATTCATATCTCATTAAGAGATTTTTCAGAGATGCTGATGATGATG 1492  
DB 841 CATGTTTACATTCATATCTCATTAAGAGATTTTTCAGAGATGCTGATGATGATGATG 900

QY 1493 ATTGATGCTGGGATATTAACAATTCATGATGATGATGATGATGATGATGATGATG 1552  
DB 901 ATTGATGCTGGGATATTAACAATTCATGATGATGATGATGATGATGATGATGATGATG 960

QY 1553 GTAGTGGAAATGAGATTATATATATTAATTAATTAATTAATTAATTAATTAATTA 1597  
DB 961 GTAGTGGAAATGAGATTATATATATTAATTAATTAATTAATTAATTAATTAATTA 1005

RESULT 3  
AX358758 1739 bp DNA linear PAT 13-FEB-2002  
LOCUS AX358758  
DEFINITION Sequence 11 from Patent W00193983.  
ACCESSION AX358758  
VERSION AX358758.1 GI:18675280  
KEYWORDS  
SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 Baker, K.P., Desnoyers, L., Gerlisen, M.E., Goddard, A.,  
Godowski, P.J., Grimaldi, J.C., Gurney, A.L., Smith, V., Stephan, J.P.,  
Watanabe, C.K. and Wood, W.I.  
Secreted and transmembrane polypeptides and nucleic acids encoding  
the same  
Patent: WO 0193983-A 11 13-DEC-2001;  
Genentech Inc. (US)  
Local/On/Qualifiers  
FEATURES  
source  
BASE COUNT 545 a 405 c 359 g 430 t  
ORIGIN  
Query Match 61.1%; Score 987; DB 6; Length 1739;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCACACTGGTGGCTGAGAAATTCGAAACCTAGTGGCCAGAGATCTATGTATG 134  
DB 3 GCACACTGGTGGCTGAGAAATTCGAAACCTAGTGGCCAGAGATCTATGTATG 62

QY 135 GGATTGAAACAATCTGTAACATAAAGATCTTAATCATGAAAAATAGATATGATTAATA 194  
DB 63 GGATTGAAACAATCTGTAACATAAAGATCTTAATCATGAAAAATAGATATGATTAATA 122

QY 195 AGTCACATATTGGCGCTGTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTA 254  
DB 123 AGTCACATATTGGCGCTGTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTA 182

QY 255 ACTCCAAAGAGTTTGGTGTCTGCTTAACATTAATCAATCTCCCATTAAGTGAACAAT 314  
DB 183 ACTCCAAAGAGTTTGGTGTCTGCTTAACATTAATCAATCTCCCATTAAGTGAACAAT 242

QY 315 CCACAAGTCTTATTTCCCTTAACACACATGATTAATTAATTAATTAATTAATTAATTA 374  
DB 243 CCACAAGTCTTATTTCCCTTAACACACATGATTAATTAATTAATTAATTAATTAATTA 302

QY 375 AACTCAGATTAAGAAATCATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 434  
DB 303 AACTCAGATTAAGAAATCATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 362

QY 435 CCCAGGTGAACACACACACACACACATGACACAGCCACCATCTCAACCTGGAG 494  
DB 363 CCCAGGTGAACACACACACACACACATGACACAGCCACCATCTCAACCTGGAG 422

QY 495 ATAGTACTGACAGGAGAACACATGCTGCAATCTGCTGAGAGTGAAGGACCACTTGGAC 554  
DB 423 ATAGTACTGACAGGAGAACACATGCTGCAATCTGCTGAGAGTGAAGGACCACTTGGAC 482

QY 555 GCAGAGCAATATGCGGAGATTTCTGAGGAGCCAGATGCTTCCAGAGGCTGATGG 614  
DB 483 GCAGAGCAATATGCGGAGATTTCTGAGGAGCCAGATGCTTCCAGAGGCTGATGG 542

QY 615 CAGGTGCTTGAAGAAAGTGAATGCTTCAACAACAGCACTTACCTGCTGCTGCTGCTG 674  
DB 543 CAGGTGCTTGAAGAAAGTGAATGCTTCAACAACAGCACTTACCTGCTGCTGCTGCTGCTG 602

QY 675 TGTTCGTGAGAGGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734  
DB 603 TGTTCGTGAGAGGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662

QY 735 CAGCTCTGAGATGCAAGAACCAAGGCTATGACAGGCTGATCTTCACTGCGCCGTTTG 794  
 DB 663 CAGCTCTGAGATGCAAGAACCAAGGCTATGACAGGCTGATCTTCACTGCGCCGTTTG 722  
 QY 795 TCATGAGCACTTCCCAAGTCACTCTGAAATGTCGCTGATCTTAAACACAATGCTGAAT 854  
 DB 723 TCATGAGCACTTCCCAAGTCACTCTGAAATGTCGCTGATCTTAAACACAATGCTGAAT 782  
 QY 855 TGTGCAAGTACTGCAACAGAGACCAAGAAAGGCTTACTGTCGTGAGGCGCTCAACACA 914  
 DB 783 TGTGCAAGTACTGCAACAGAGACCAAGAAAGGCTTACTGTCGTGAGGCGCTCAACACA 842  
 QY 915 TGGCTGTGCTGCACTGATCATATGATTTTAAAGAAAGAAAGTTTCTTATCTTAGCA 974  
 DB 843 TGGCTGTGCTGCACTGATCATATGATTTTAAAGAAAGAAAGTTTCTTATCTTAGCA 902  
 QY 975 AACAGAAAAGAGCCTTTTAAAGGTCAATGTGGGTAGATGATTTGAAAAATTTCA 1034  
 DB 903 AACAGAAAAGAGCCTTTTAAAGGTCAATGTGGGTAGATGATTTGAAAAATTTCA 962  
 QY 1035 ATACATTAGTGTCTCCAAATGCAACA 1061  
 DB 963 ATACATTAGTGTCTCCAAATGCAACA 989  
 RESULT 4  
 AX362251 1739 bp DNA linear PAT 15-FEB-2002  
 LOCUS Sequence 11 from Patent WO0208288.  
 DEFINITION AX362251  
 ACCESSION AX362251  
 VERSION AX362251.1 GI:18694583  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Baker, K.P., Desnoyers, L., Gerritsen, M.E., Goddard, A.,  
 Godwin, P.J., Grimaldi, J.C., Gurney, A.L., Smith, V., Stephan, J.P.,  
 Malanbe, C.K. and Wood, W.F.,  
 Title Secreted and transmembrane polypeptides and nucleic acids encoding  
 the same  
 Journal Patent: WO 0208288-A 11 31-JAN-2002;  
 Genentech, Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1. 1138  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 545 a 405 c 359 g 430 t  
 ORIGIN  
 Query Match 61.1%; Score 987; DB 6; Length 1739;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 CCACAAAGTCTTATTCCTTAAACACAGCTGATATGATTTAAAGCCACTAACAGAGACTG 302  
 DB 375 AACTGAGATTAAGAAATATGAGAAATGATGACAGATTCCTGACGAGCTTTTCA 414  
 QY 303 AACTGAGATTAAGAAATATGAGAAATGATGACAGATTCCTGACGAGCTTTTCA 362  
 DB 435 CCGACGTGAACACCAACGACGCGCCACATAGACAGCCACATCTCTCAACCTGAG 494  
 QY 363 CCGACGTGAACACCAACGACGCGCCACATAGACAGCCACATCTCTCAACCTGAG 422  
 DB 495 ATACGTACTGAGGAGGAGACCACTGACATCTGCTGAGAGGTGAGGACACTTGGAC 554  
 QY 423 ATACGTACTGAGGAGGAGACCACTGACATCTGCTGAGAGGTGAGGACACTTGGAC 482  
 DB 555 GCGAGAACCATATGCGGCGGATTTCTGAGAGGCAAGATGCTTCCGACGCTATG 614  
 QY 483 GCGAGAACCATATGCGGCGGATTTCTGAGAGGCAAGATGCTTCCGACGCTATG 542  
 DB 615 CAGGTCTTTCAGAAAGTGTGACTGACTTCAACACAGGACCTTACTGCTGAGTCTGACTG 674  
 DB 543 CAGGTCTTTCAGAAAGTGTGACTGACTTCAACACAGGACCTTACTGCTGAGTCTGACTG 602  
 QY 675 TGTGTCGAGAGGCGCAAGTCTCTGCTGCTGCTCTCATCCACCCAGTCAAGCGATG 734  
 DB 603 TGTGTCGAGAGGCGCAAGTCTCTGCTGCTGCTCTCATCCACCCAGTCAAGCGATG 662  
 QY 735 CAGCTCTGAGATGCAAGAACCAAGGCTATGACAGGCTGATTTTAAAGCAATGCTGAAT 794  
 DB 663 CAGCTCTGAGATGCAAGAACCAAGGCTATGACAGGCTGATTTTAAAGCAATGCTGAAT 722  
 QY 795 TCAATGGCACTGCGCAAGTCTGCTGATGATGCTGCTGATGCTTAAACACAATGCTGAAT 854  
 DB 723 TCAATGGCACTGCGCAAGTCTGCTGATGATGCTGCTGATGCTTAAACACAATGCTGAAT 782  
 QY 855 TGTGCAAGTACTGCAACAGAGACCAAGAAAGGCTTACTGTCGTGAGGCGCTCAACACA 914  
 DB 783 TGTGCAAGTACTGCAACAGAGACCAAGAAAGGCTTACTGTCGTGAGGCGCTCAACACA 842  
 QY 915 TGGCTGTGCTGCACTGATCATATGATTTTAAAGAAAGAAAGTTTCTTATCTTAGCA 974  
 DB 843 TGGCTGTGCTGCACTGATCATATGATTTTAAAGAAAGAAAGTTTCTTATCTTAGCA 902  
 QY 975 AACAGAAAAGAGCCTTTTAAAGGTCAATGTGGGTAGATGATTTGAAAAATTTCA 1034  
 DB 903 AACAGAAAAGAGCCTTTTAAAGGTCAATGTGGGTAGATGATTTGAAAAATTTCA 962  
 QY 1035 ATACATTAGTGTCTCCAAATGCAACA 1061  
 DB 963 ATACATTAGTGTCTCCAAATGCAACA 989  
 RESULT 5  
 AX083510 1138 bp DNA linear PAT 28-FEB-2001  
 LOCUS Sequence 52 from Patent WO0112662.  
 DEFINITION AX083510  
 ACCESSION AX083510  
 VERSION AX083510.1 GI:13185320  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1138)  
 Authors Lal, P., Yue, H., Tang, Y.T., Bandman, O., Burford, N., Azimzal, Y.,  
 Baughn, M.R., Lu, D.A. and Patterson, C.  
 Title Membrane associated proteins  
 Journal Patent: WO 0112662-A 52 22-FEB-2001;  
 Incyte Genomics, Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1. 1138  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"



BASE COUNT 334 a 274 c 241 g 289 t  
 ORIGIN

Query Match 60.8%; Score 983; DH 6; Length 1138;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CAACCTGCACTTATATCTGCAATTTATTTGGTATAGACAGAGATATGCCAGTACAC 78  
 DB 1 CAACCTGCACTTATATCTGCAATTTATTTGGTATAGACAGAGATATGCCAGTACAC 78  
 QY 79 ACTGCTGCTTCAGAGAAATTCACACCTAGCTGGCAGAGGCTATATGTATGGAT 138  
 DB 61 ACTGCTGCTTCAGAGAAATTCACACCTAGCTGGCAGAGGCTATATGTATGGAT 120  
 QY 139 TGAACAATCTGTAACTAAAGATCTTATCATGAAATTAAGTATGATTAATTAAGTC 198  
 DB 121 TGAACAATCTGTAACTAAAGATCTTATCATGAAATTAAGTATGATTAATTAAGTC 180  
 QY 199 ACTATTCGACCTGTTTATATATAGCTCCTGATCATTTTACAGTTTCCAGAACTC 258  
 DB 181 ACTATTCGACCTGTTTATATATAGCTCCTGATCATTTTACAGTTTCCAGAACTC 240  
 QY 459 GACAAAGTTTGGTCTGCTTAATTTATCATCTCCCTCATTAAGTGAACAATCCAC 318  
 DB 241 GACAAAGTTTGGTCTGCTTAATTTATCATCTCCCTCATTAAGTGAACAATCCAC 300  
 QY 319 AAGTCTTATTCCTTAACACACGATGATATCATTAAGCAGCAGACAGAGTCACT 378  
 DB 301 AAGTCTTATTCCTTAACACACGATGATATCATTAAGCAGCAGACAGAGTCACT 360  
 QY 379 CAGATTAAGCAATGATAGAAATGATATGATGATGATGATGATGATGATGATGAT 438  
 DB 361 CAGATTAAGCAATGATAGAAATGATATGATGATGATGATGATGATGATGATGAT 420  
 QY 439 GGTGAC 498  
 DB 421 GGTGAC 480  
 QY 499 GTACTGACGAGGAGACACACACACACACACACACACACACACACACACACAC 558  
 DB 481 GTACTGACGAGGAGACACACACACACACACACACACACACACACACACACAC 540  
 QY 559 GAGCAATATGCGCGGATTTCTGAGCGGACGATGTTCCCGACCGCTATGTCAGG 618  
 DB 541 GAGCAATATGCGCGGATTTCTGAGCGGACGATGTTCCCGACCGCTATGTCAGG 600  
 QY 619 TGTTCAGGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 678  
 DB 601 TGTTCAGGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 660  
 QY 679 GTGGAGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738  
 DB 661 GTGGAGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 QY 739 TCTCTGAGTGAAGCAACAGCTATGACAGGATATGCTGCTGCTGCTGCTGCTG 798  
 DB 721 TCTCTGAGTGAAGCAACAGCTATGACAGGATATGCTGCTGCTGCTGCTGCTG 780  
 QY 799 TGGACTTCCCAAGTCAAGCTGATATGATGATGATGATGATGATGATGATGAT 858  
 DB 781 TGGACTTCCCAAGTCAAGCTGATATGATGATGATGATGATGATGATGATGAT 840  
 QY 859 GCAGTACTGCAACAGAGACCAAGAGCTTACTGCTGCTGCTGCTGCTGCTGCTG 918  
 DB 841 GCAGTACTGCAACAGAGACCAAGAGCTTACTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 919 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978  
 DB 901 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 979 AGAAGAGAGCTCTTTGAAAGGT 1001

DB 961 AGAAGAGAGCTCTTTGAAAGGT 983

RESULT 6  
 LOCUS AP000450/c  
 DEFINITION Homo sapiens chromosome 11 clone XXP1-307 map 11q23, WORKING DRAFT  
 ACCESSION AP000450  
 VERSION AP000450.3 GI:8118763  
 KEYWORDS HTGS-PHASE1; HTGS-DRAFT.  
 SOURCE Homo sapiens DNA, clone: XXP1-307.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 JOURNAL Homo sapiens 81,975 genomic DNA of 11q23  
 REFERENCE Published Only in Database (1999)  
 AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-SEP-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923, Fax: 81-42-778-9924)  
 COMMENT On May 31, 2000 this sequence version replaced gi:7630227.

Center: RIKEN Genomic Sciences Center (GSC)  
 Center code: RIKEN  
 Web site: http://hgp.gsc.riken.go.jp/  
 Contact: hattori@gsc.riken.go.jp  
 Project Information  
 Center project name: HumDrafc11  
 Center clone name: XXP1-307

Summary Statistics  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 79702 bases at least Q40  
 Consensus quality: 80980 bases at least Q30  
 Consensus quality: 81428 bases at least Q20  
 Insert size: 81575; sum-of-coverage  
 Quality coverage: 5.78x in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 35184 contig of 35184 bp in length  
 35285 51779 contig of 16495 bp in length  
 51880 63010 contig of 11131 bp in length  
 63111 73886 contig of 10776 bp in length  
 73987 81975 contig of 7989 bp in length  
 Sequence updated (18-Apr-2000)

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
 1 35184: contig of 35184 bp in length



30): an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Teleno, M., Caranese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

#### VECTOR: pBAC3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-720G3; the clone sequenced to the right is RP11-643E23. Actual start of this clone is at base position 1 of RP11-172C16; actual end is at base position 146841 of RP11-172C16.

The region from base position 130951 to 131438 represents the best possible assembly of an imperfect TA repeat. Restriction digest information from EcoRV, HindIII, and BamHI indicates that approximately 150 to 250 bases may be missing from the final sequence.

#### FEATURES

##### source

1..146841  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone\_jib="RPC1-11"  
418..575  
/rpt\_family="MER1\_type"  
774..946  
/rpt\_family="L1"  
1114..1313  
/rpt\_family="L1"  
1365..1452  
/rpt\_family="ERV"  
1425..1510  
/rpt\_family="ERV"  
1541..2262  
/rpt\_family="L1"  
2269..2957  
/rpt\_family="L1"  
8426..8655  
/rpt\_family="L1"  
8656..8957  
/rpt\_family="Alu"  
8958..8996  
/rpt\_family="L1"  
9029..9453  
/rpt\_family="L1"  
10059..10457  
/rpt\_family="MALR"  
10468..10610  
/rpt\_family="L1"  
10659..10807  
/rpt\_family="L2"  
10843..11005  
/rpt\_family="ERV1"  
12544..12725  
/rpt\_family="L1"  
12830..13126  
repeat\_region

/rpt\_family="Alu"  
13433..13674  
/rpt\_family="L1"  
14419..14587  
/rpt\_family="MIR"  
15199..15545  
/rpt\_family="L2"  
16012..16077  
/rpt\_family="L2"  
17372..17741  
/note="similar to  
17555..18180  
/note="similar to  
17696..18060  
/note="similar to  
17999..18157  
/rpt\_family="L2"  
18220..18324  
/rpt\_family="L1"  
18325..18401  
/rpt\_family="Alu"  
18408..18538  
/rpt\_family="Alu"  
18539..19181  
/rpt\_family="L1"  
19454..19661  
/rpt\_family="L1"  
19705..20140  
/rpt\_family="L1"  
20162..20624  
/note="similar to  
20371..20690  
/note="similar to  
22353..22774  
/rpt\_family="MALR"  
23334..23715  
/rpt\_family="ERV1"  
23738..23802  
/rpt\_family="MALR"  
23803..24380  
/rpt\_family="ERV1"  
24381..24774  
/rpt\_family="MALR"  
24925..25100  
/rpt\_family="L2"  
25169..25250  
/rpt\_family="MER53"  
25524..25566  
/rpt\_family="L2"  
25567..25668  
/rpt\_family="U13"  
26564..26944  
/rpt\_family="L2"  
26947..27061  
/rpt\_family="L2"  
28123..28317  
/rpt\_family="L2"  
29799..30188  
/rpt\_family="L2"  
30312..30484  
/rpt\_family="MER1\_type"  
38493..38692  
/rpt\_family="Alu"  
38712..38863  
/rpt\_family="L1"  
39064..39272  
/rpt\_family="L1"  
39323..39868  
repeat\_region

#### Query Match

Best Local Similarity 45.78; Score 738; DB 9; Length 146841;  
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 264 AGTTTGCTGCTGCTAAACTATCATCTCCCTCATTTACTGGAACCACTCCAAAGT 323
    |||||||
Db 81461 AGTTTGCTGCTGCTAAACTATCATCTCCCTCATTTACTGGAACCACTCCAAAGT 81402

QY 324 CTTATTCCCTAAACACCACTGATATCATTTAAAGCCACTAGAGAGACTGAACAGAA 383
    |||||||
Db 81401 CTTATTCCCTAAACACCACTGATATCATTTAAAGCCACTAGAGAGACTGAACAGAA 81342

QY 384 TAAAGAAATCATAGAGAAATCATAGATACAGATCCCAAGCCAGACCTTCCACCCAGTGA 443
    |||||||
Db 81341 TAAAGAAATCATAGAGAAATCATAGATACAGATCCCAAGCCAGACCTTCCACCCAGTGA 81282

QY 444 ACACCAACACAGCCCAACATAGACAGCCACCATCTTAAACCTGAGATAGACTACT 503
    |||||||
Db 81281 ACACCAACACAGCCCAACATAGACAGCCACCATCTTAAACCTGAGATAGACTACT 81222

QY 504 GCAGGGGAGACCAAGCTGACATCCCTGCTGAGAGGTAGAGGACCACTTGGAGCCAGGAAAGC 563
    |||||||
Db 81221 GCAGGGGAGACCAAGCTGACATCCCTGCTGAGAGGTAGAGGACCACTTGGAGCCAGGAAAGC 81162

QY 564 AATATGCGGGGATTTCTCTGAGGCGCAGATGTTCCCAAGCGCTGATGGCAGTGCTT 623
    |||||||
Db 81161 AATATGCGGGGATTTCTCTGAGGCGCAGATGTTCCCAAGCGCTGATGGCAGTGCTT 81102

QY 624 CAGGAAAGGTGACTGACTTCAACAGGACCACTGCTGACCTTCACTGCTTCTGGC 683
    |||||||
Db 81101 CAGGAAAGGTGACTGACTTCAACAGGACCACTGCTGACCTTCACTGCTTCTGGC 81042

QY 684 AGGGCCAGGCTCTGCTGCTGCTGCTGCTATGACACCCAGTGAAGGGGGTGCACCTCT 743
    |||||||
Db 81041 AGGGCCAGGCTCTGCTGCTGCTGCTGCTATGACACCCAGTGAAGGGGGTGCACCTCT 80982

QY 744 GGAGTGCAGAGAACCAAGGCTATGACAGGGTATCTTCACTGGCCAGTTTGTCAATGGCA 803
    |||||||
Db 80981 GGAGTGCAGAGAACCAAGGCTATGACAGGGTATCTTCACTGGCCAGTTTGTCAATGGCA 80922

QY 804 CTTCCCAAGTCACTGCTGATGAGGCTGATCTTAAACACAATGCTGAATTTGGCCAGT 863
    |||||||
Db 80921 CTTCCCAAGTCACTGCTGATGAGGCTGATCTTAAACACAATGCTGAATTTGGCCAGT 80862

QY 864 ACCTGGACAAAGAGACCAAGAAAGGCTTACTGTGAGGCGCTCAACATGCGCTGTG 923
    |||||||
Db 80861 ACCTGGACAAAGAGACCAAGAAAGGCTTACTGTGAGGCGCTCAACATGCGCTGTG 80802

QY 924 CTGCACTCACTCACTATTTCTTAAGACAGAAAGTTTCTTATCTTGGCAACAAAGAA 983
    |||||||
Db 80801 CTGCACTCACTCACTATTTCTTAAGACAGAAAGTTTCTTATCTTGGCAACAAAGAA 80742

QY 984 AGAGCTCTTTGAAAGGT 1001
    |||||||
Db 80741 AGAGCTCTTTGAAAGGT 80724

RESULT 8
AC068041 187852 bp DNA linear HTG 02-SRP-2000
LOCUS AC068041
DEFINITION Homo sapiens chromosome 11 clone RP11-643E23, WORKING DRAFT
ACCESSION AC068041.4 GI:965027
VERSION AC068041.4
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 187852)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187852)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Genome Sequencing Center, Washington

```

## COMMENT

University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Sep 2, 2000 this sequence version replaced g1:8469045.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Project information  
Center project name: H\_NH0643E23  
Summary Statistics

Sequencing vector: p13; 100%  
Sequencing vector: plasmid; 0%  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 183822 bases at least Q43  
Consensus quality: 185080 bases at least Q10  
Consensus quality: 185802 bases at least Q20  
Insert size: 186000; agarose-fp  
Insert size: 187052; sum-of-coverage  
Quality coverage: 7.61 in Q20 bases; sum-of-coverage  
Quality coverage: 7.02 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1 1761: contig of 1761 bp in length
* 1762: gap of unknown length
* 1862: contig of 6473 bp in length
* 8335: gap of unknown length
* 8435: contig of 7479 bp in length
* 15914: gap of unknown length
* 16014: contig of 7089 bp in length
* 23103: gap of unknown length
* 23202: gap of unknown length
* 31255: contig of 7953 bp in length
* 31256: gap of unknown length
* 31256: gap of unknown length
* 45995: contig of 14740 bp in length
* 46096: gap of unknown length
* 77307: contig of 51211 bp in length
* 77406: gap of unknown length
* 144505: contig of 67099 bp in length
* 144506: gap of unknown length
* 144606: contig of 43247 bp in length.
Location/Qualifiers
1.187852
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-643E23"
1.1761
/note="assembly_name:Contig34"
1862.8334
/note="assembly_name:Contig35"
8435.15913
/note="assembly_name:Contig36"
16014.23102
/note="assembly_name:Contig37"
23203.31155
/note="assembly_name:Contig38"
31256.45995
/note="assembly_name:Contig39"
46096.77306
/note="assembly_name:Contig40"
vector_end:SP6
vector_slide:right"
77407.144505
/note="assembly_name:Contig42

```

## FEATURES

source

```

misc_feature
    /db_xref="taxon:9606"
    /clone="RP11-643E23"
    1.1761
/note="assembly_name:Contig34"
1862.8334
/note="assembly_name:Contig35"
8435.15913
/note="assembly_name:Contig36"
16014.23102
/note="assembly_name:Contig37"
23203.31155
/note="assembly_name:Contig38"
31256.45995
/note="assembly_name:Contig39"
46096.77306
/note="assembly_name:Contig40"
vector_end:SP6
vector_slide:right"
77407.144505
/note="assembly_name:Contig42

```

clone\_end:77  
vector\_side:right\*  
misc\_feature /note="assembly\_name:Contig11"  
BASE COUNT 55964 a 35742 c 36408 g 58936 t 802 others  
ORIGIN

Query Match 45.7%; Score 738; DB 2; Length 187852;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Y 264 AGCTTGTCTGCTCTAACTATTCATTCCTCCATTCAGCAACCTCCACAACCT 323
Db 64776 AGCTTGTCTGCTCTAACTATTCATTCCTCCATTCAGCAACCTCCACAACCT 64835

Y 324 CTTATTCCCTAAACACACATGATATCATTTAAAGCAGCTAAAGAGTGAAGTGA 383
Db 64816 CTTATTCCCTAAACACACATGATATCATTTAAAGCAGCTAAAGAGTGAAGTGA 64895

Y 384 TAAAGAAATCATAGAGAAAGTATGATGAGATGCTCCACCCAGACCTTTACCCAGTGA 443
Db 64896 TAAAGAAATCATAGAGAAAGTATGATGAGATGCTCCACCCAGACCTTTACCCAGTGA 64955

Y 444 ACACCCACCAACCCGACACATAGACACACCCACATCCCTGACGATAGTACT 503
Db 64956 ACACCCACCAACCCGACACATAGACACACCCACATCCCTGACGATAGTACT 65015

Y 504 GCAGGGGAGACACGATGACATCTGCTGGAGGTGAGGAGCAGTTGGAGCGAGAGAAC 563
Db 65016 GCAGGGGAGACACGATGACATCTGCTGGAGGTGAGGAGCAGTTGGAGCGAGAGAAC 65075

Y 564 AATATGCGCGGATTTCTCTGAGGGCCAGCATGCTCTCCAGCGCTGATGGAGGTGCTT 623
Db 65076 AATATGCGCGGATTTCTCTGAGGGCCAGCATGCTCTCCAGCGCTGATGGAGGTGCTT 65135

Y 624 CAGCAAGGTGACTGACTTCAACACCGACACCTGCTGAGTTGACCTGCTGCTGG 683
Db 65136 CAGCAAGGTGACTGACTTCAACACCGACACCTGCTGAGTTGACCTGCTGCTGG 65195

Y 684 AGGGTCAAGTCTCTGCTGCTGCTGCTATCCACCCAGTGAAGGGGTGACGCTCTCT 743
Db 65196 AGGGTCAAGTCTCTGCTGCTGCTGCTATCCACCCAGTGAAGGGGTGACGCTCTCT 65255

Y 744 GAGTGTCAAGAACCAAGCCTATGACAGGGTATGCTTCACTGGCCAGTTTGCATGACGA 803
Db 65256 GAGTGTCAAGAACCAAGCCTATGACAGGGTATGCTTCACTGGCCAGTTTGCATGACGA 65315

Y 804 CTTCGCCAGTCTGACTGCTGAATGTGCGCTGATCTTAACACAAATGCTGAATTTGCCAGT 863
Db 65316 CTTCGCCAGTCTGACTGCTGAATGTGCGCTGATCTTAACACAAATGCTGAATTTGCCAGT 65375

Y 864 ACCTGGACACAGAGACCAAGAGGTTTCTACTGTGTAGGCGCTCAACACATGCGCTGTG 923
Db 65376 ACCTGGACACAGAGACCAAGAGGTTTCTACTGTGTAGGCGCTCAACACATGCGCTGTG 65435

Y 924 CTGCAGTCTGACTGACTGATTTCAAGCAAGAAATTTCTTATTTAGCAACAAAGAAA 983
Db 65436 CTGCAGTCTGACTGACTGATTTCAAGCAAGAAATTTCTTATTTAGCAACAAAGAAA 65495

Y 984 AGAGCCTCTTTGAAAGGT 1001
Db 65496 AGAGCCTCTTTGAAAGGT 65513
```

ESOLUT 9  
X45H380/C 621 bp DNA linear PAT 08-JUL-2002  
FEINITION Sequence 5 from Patent WO0246411.  
ERSTON AX45H380  
EYWRDS AX45H380.1 GI:21725048  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
JOURNAL Lasek, A.W., Krasnow, R. and Baughn, M.R.  
Patent: WO 0246411-A 5 13-JUN-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1.621  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 71951705V1"

BASE COUNT 142 a 138 c 179 g 162 t  
ORIGIN

Query Match 34.5%; Score 558; DB 6; Length 621;  
Best Local Similarity 99.8%; Pred. No. 1; 46-305;  
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Y 210 TGTGTTTATATTAGGCTCTGATCATTTTTCAGATTTCAGAACTCCAGAAAGTTT 269
Db 621 TGTGTTTATATTAGGCTCTGATCATTTTTCAGATTTCAGAACTCCAGAAAGTTT 562

Y 270 GGTGCTCTAAACTTATTCATCTCCCTCATTTACTGGAACAACTCCAGAAAGTTTAT 329
Db 561 GGTGCTCTAAACTTATTCATCTCCCTCATTTACTGGAACAACTCCAGAAAGTTTAT 502

Y 330 TCCCTAAACACCATGATATCATTAAGCCATTAAGAGATGAACTCAGATTAAGG 389
Db 501 TCCCTAAACACCATGATATCATTAAGCCATTAAGAGATGAACTCAGATTAAGG 442

Y 390 AAATCATAGAGAACTAGATGAGAGATCCACCAAGCCTTTACCCAGCTGAACACA 449
Db 441 AAATCATAGAGAACTAGATGAGAGATCCACCAAGCCTTTACCCAGCTGAACACA 382

Y 450 CCACCAAGCCAGACATAGACACACCCACATCCCTGACAGATAGCTACTGCAAGG 509
Db 381 CCACCAAGCCAGACATAGACACACCCACATCCCTGACAGATAGCTACTGCAAGG 322

Y 510 GAGACAGCTGCATCTGCTGAGGTGAGGAGAACCTTTGGAGCCAGAGAAATATG 569
Db 321 GAGACAGCTGCATCTGCTGAGGTGAGGAGAACCTTTGGAGCCAGAGAAATATG 262

Y 570 GCGGGGATTTCTGAGGGCCAGAGATCTTCCCGAGCCTGATAGCAGGTGCTTCAAGAA 629
Db 261 GCGGGGATTTCTGAGGGCCAGAGATCTTCCCGAGCCTGATAGCAGGTGCTTCAAGAA 202

Y 630 AGTGAAGTACTTCAACACAGGGACCTACCTGCTGACGTTCACTGTTCTGGAGGGCC 689
Db 201 AGTGAAGTACTTCAACACAGGGACCTACCTGCTGACGTTCACTGTTCTGGAGGGCC 142

Y 690 AGTCTCTCTGCTGCTGCTGCTATCCACCCAGTGAAGGGGTGTCAGTCTGAGATG 749
Db 141 AGTCTCTCTGCTGCTGCTGCTATCCACCCAGTGAAGGGGTGTCAGTCTGAGATG 82

Y 750 CAAGGAACCAAGCCTATGACAGGGGTGATTCACCTGCGAGTTTGCATAGGCACTTCC 809
Db 81 CAAGGAACCAAGCCTATGACAGGGGTGATTCACCTGCGAGTTTGCATAGGCACTTCC 22

Y 810 AAGTCACT 818
Db 21 AAGTCACT 13
```

RESULT 10  
AK000134 1704 bp mRNA linear PRI 22-FEB-2000  
LOCUS Homo sapiens cDNA FLJ20127 f1s, clone COL06176.  
DEFINITION AK000134  
ACCESSION AK000134  
VERSION AK000134.1 GI:7020022  
KEYWORDS Oligo capping; f1s (full insert sequence).  
SOURCE Homo sapiens colon cDNA to mRNA, clone\_11b; COL clone:COL06176.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

**REFERENCE**  
**AUTHORS** Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.

**TITLE** NEDO human cDNA sequencing project

**JOURNAL** Unpublished

**REFERENCE** 2 (bases 1 to 1704)

**AUTHORS** Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.

**TITLE** Direct Submission

**JOURNAL** Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail:cdna@iis.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

**COMMENT** NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'-63'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

**FEATURES**  
**Source** 1..1704  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="COL06176"  
 /tissue-type="colon"  
 /clone\_lib="COL"  
 /note="cloning vector pWE18SFL3"  
 558..1040  
 /note="unnamed protein product"  
 /protein\_id="BAA90965.1"  
 /db\_xref="GI:7020023"  
 /translation="MEKFNITISAKCNKEVAMKEKFGKSTIPSGHVMNTWNPY  
 SCSLATKMECELRGLIYLMGDSIIOMWEYFKASITLKSVDLESKLOHOLAVD  
 LDRNINQMCKYCPILGISMTYSYKMEVLTALIDRIGKEKKYCHCFPGAPQLSH  
 \*"

**CDS**  
 558..1040  
 /note="unnamed protein product"  
 /protein\_id="BAA90965.1"  
 /db\_xref="GI:7020023"  
 /translation="MEKFNITISAKCNKEVAMKEKFGKSTIPSGHVMNTWNPY  
 SCSLATKMECELRGLIYLMGDSIIOMWEYFKASITLKSVDLESKLOHOLAVD  
 LDRNINQMCKYCPILGISMTYSYKMEVLTALIDRIGKEKKYCHCFPGAPQLSH  
 \*"

**BASE COUNT** 597 a 366 c 308 g 433 t

**ORIGIN**

Query Match 32.7%; Score 528; DB 9; Length 1704;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-288;  
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCCATGCTAAGGTAACAACTGCACTATATCGCAATTTATTTGGATAGCAA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1 ATCCATGCTAAGGTAACAACTGCACTATATCGCAATTTATTTGGATAGCAA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 61 GAGGTATGCCAGTACGACACTGTGGCTTCAGAGAAATTCAGACCTAGCTGGCAG 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 61 GAGGTATGCCAGTACGACACTGTGGCTTCAGAGAAATTCAGACCTAGCTGGCAG 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 121 AGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 121 AGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 181 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 181 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 241 TACAGTTTCAGAACTGCACAAAGGTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 241 TACAGTTTCAGAACTGCACAAAGGTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 301 TTACTGGACAACTGCACAAAGGTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 301 TTACTGGACAACTGCACAAAGGTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 361 ACTAACGAGAGTGAAGTCAAGATTAAGAAATCATAGAAATCATAGAAATCATAGAAATCAT 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 361 ACTAACGAGAGTGAAGTCAAGATTAAGAAATCATAGAAATCATAGAAATCATAGAAATCAT 420

Qy 421 ACCGACACCTTTACACCTGTAACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 421 ACCGACACCTTTACACCTGTAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 481 CCTCAACCCCTGAGATAGTACTGTCAGAGGAGGAGACCAACCTGCAATCTCT 528  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 481 CCTCAACCCCTGAGATAGTACTGTCAGAGGAGGAGACCAACCTGCAATCTCT 528  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

**RESULT 11**  
 AX458381/c 545 bp DNA Linear PAT. Ch. INTL. 277  
 LOCUS Sequence 6 from Patent WO0246411.  
 ACCESSION AX458381  
 VERSION AX458381.1 GI:21725049  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

**REFERENCE**  
**AUTHORS** Lasex, A.W., Krasnow, R. and Baughn, M.R.  
**TITLE** Intestinal proteins  
**JOURNAL** Patent: WO 0246411-A 6 13-JUN-2002;  
 Incyte Genomics, Inc. (US)

**FEATURES**  
**Source** 1..545  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 70255975V1"  
 110 a 129 c 161 g 145 t

**BASE COUNT** 110 a 129 c 161 g 145 t

**ORIGIN**

Query Match 25.3%; Score 409; DB 6; Length 545;  
 Best Local Similarity 99.6%; Pred. No. 9.9e-221;  
 Matches 509; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

319 AAGGTCTTATTCCTTAAACACCACTGATATCTTAAACCACTTAACAGACTGAACT 378  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 511 AAGGTCTTATTCCTTAAACACCACTGATATCTTAAACCACTTAACAGACTGAACT 452  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 379 CAGATTAAGCAATCATAGAAATCATAGAAATCATAGAAATCATAGAAATCATAGAAATCAT 448  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 451 CAGATTAAGCAATCATAGAAATCATAGAAATCATAGAAATCATAGAAATCATAGAAATCAT 492  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 439 CGTGAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 391 CGTGAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 392  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 499 GTACTGCAGGAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 558  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 331 GTACTGCAGGAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 272  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 559 GAAGCAATATGCGGGGATTTCTGTAGGGCCAGGATGTCTTCCCAAGCGCTGATGATGATG 648  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 271 GAAGCAATATGCGGGGATTTCTGTAGGGCCAGGATGTCTTCCCAAGCGCTGATGATGATG 212  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 619 TGCTTCAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 678  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 211 TGCTTCAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 152  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 679 CTGGAGGGGCGGAGGCTCTGT 738  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 151 CTGGAGGGGCGGAGGCTCTGT 92  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 739 TCTCTGAGATGCAAGGAACCAAGGCTATGACAGGATGATCTTACTGGCGAGTTTGTCAA 798  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 91 TCTCTGAGATGCAAGGAACCAAGGCTATGACAGGATGATCTTACTGGCGAGTTTGTCAA 32  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 799 TGGCACTTCCCAAGTCACTCTGAATGAGGC 829  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||



TITLE  
JOURNAL  
COMMENT

McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Menzies, L., Minova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Olliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigg, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 7, 2000 this sequence version replaced g1:7651863.  
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: 19810

Center clone name: 720.G.3

Summary Statistics

Sequencing vector: M13: W77815, 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 161804 bases at least Q40

Consensus quality: 170435 bases at least Q30

Consensus quality: 173827 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 175394; sum-of-contigs

Quality coverage: 3.6 in Q20 bases; agarose-fp

Quality coverage: 3.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 32 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1247: contig of 1247 bp in length

1248 1347: gap of 100 bp

1348 2922: contig of 1575 bp in length

2923 3022: gap of 100 bp

3023 4326: contig of 1304 bp in length

4327 4426: gap of 100 bp

4427 5770: contig of 1344 bp in length

5771 5870: gap of 100 bp

5871 7537: contig of 1667 bp in length

7538 7637: gap of 100 bp

7638 10278: contig of 2642 bp in length

10279 10378: gap of 100 bp

10379 12229: contig of 1851 bp in length

12230 12329: gap of 100 bp

12330 14202: contig of 1873 bp in length

14203 14302: gap of 100 bp

14303 17175: contig of 2873 bp in length

17176 17275: gap of 100 bp

17276 20747: contig of 3472 bp in length

20748 20847: gap of 100 bp

20848 24235: contig of 3388 bp in length

24236 24335: gap of 100 bp

24336 28438: contig of 4103 bp in length

28439 28538: gap of 100 bp

28539 32248: contig of 3710 bp in length

32249 32348: gap of 100 bp

32349 35121: contig of 2773 bp in length

35122 35221: gap of 100 bp

FEATURES

source

35222 38838: contig of 3617 bp in length  
38839 38938: gap of 100 bp  
38939 43464: contig of 4526 bp in length  
43465 43564: gap of 100 bp  
43565 46133: contig of 4569 bp in length  
46134 48233: gap of 100 bp  
48234 52993: contig of 4760 bp in length  
52994 53093: gap of 100 bp  
53094 57397: contig of 4304 bp in length  
57398 57497: gap of 100 bp  
57498 62668: contig of 5171 bp in length  
62669 62768: gap of 100 bp  
62769 67716: contig of 4948 bp in length  
67717 67816: gap of 100 bp  
67817 74417: contig of 6601 bp in length  
74418 74517: gap of 100 bp  
74518 82397: contig of 7880 bp in length  
82398 82497: gap of 100 bp  
82498 89948: contig of 7451 bp in length  
89949 90048: gap of 100 bp  
90049 98408: contig of 8360 bp in length  
98409 98508: gap of 100 bp  
98509 108229: contig of 9721 bp in length  
108230 108329: gap of 100 bp  
108330 118955: contig of 10626 bp in length  
118956 119055: gap of 100 bp  
119056 129091: contig of 10036 bp in length  
129092 129191: gap of 100 bp  
129192 140666: contig of 11475 bp in length  
140667 140765: gap of 100 bp  
140766 150098: contig of 9332 bp in length  
150099 150198: gap of 100 bp  
150199 161240: contig of 11042 bp in length  
161241 161340: gap of 100 bp  
161341 178494: contig of 17154 bp in length.

Location/Qualifiers

1. 178494  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-72063"  
/clone\_11b="RPC1-11 Human Male BAC"  
1. 1247  
/note="assembly-fragment"  
1348..2922  
/note="assembly-fragment"  
3023..4326  
/note="assembly-fragment"  
4427..5770  
/note="assembly-fragment"  
5871..7537  
/note="assembly-fragment"  
7638..10278  
/note="assembly-fragment"  
10379..12229  
/note="assembly-fragment"  
12330..14202  
/note="assembly-fragment"  
vector\_side:left  
14303..17175  
/note="assembly-fragment"  
17276..20747  
/note="assembly-fragment"  
20848..24235  
/note="assembly-fragment"  
24336..28438  
/note="assembly-fragment"  
28539..32248  
/note="assembly-fragment"  
32349..35121  
/note="assembly-fragment"



```
misc_feature 35222..38838
/note="assembly-fragment"
misc_feature 38939..43464
/note="assembly-fragment"
misc_feature 43565..48133
/note="assembly-fragment"
misc_feature 48234..52993
/note="assembly-fragment"
misc_feature 53094..57397
/note="assembly-fragment"
misc_feature 57498..62668
/note="assembly-fragment"
misc_feature 62769..67716
/note="assembly-fragment"
misc_feature 67817..74417
/note="assembly-fragment"
misc_feature 74518..82397
/note="assembly-fragment"
misc_feature 82498..89948
/note="assembly-fragment"
misc_feature 90049..98408
/note="assembly-fragment"
misc_feature 98509..108229
/note="assembly-fragment"
misc_feature 108330..118955
/note="assembly-fragment"
misc_feature 119056..129091
/note="assembly-fragment"
```

Query Match 21.7% Score 350; DB 2; Length 178494;  
Best Local Similarity 99.8%; Pred.No. 4.3e-107;  
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1062 CACTGAAGTCAGTCGATCTGCATGATCTGGAATAATTCACACACACCTTCTGATTT 1121
|||||
DB 150599 CACTGAAGTCAGTCGATCTGCATGATCTGGAATAATTCACACACACCTTCTGATTT 150540

QY 1122 TCGATGGAACATCAACATCCAGTGGCAAAATATTGTATCCCTGATAGATCAATGA 1181
|||||
DB 150539 TCGATGGAACATCAACATCCAGTGGCAAAATATTGTATCCCTGATAGATCAATGA 150480

QY 1182 CCAATTCAGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1241
|||||
DB 150479 CCAATTCAGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 150420

QY 1242 AAAATCTGTCATCTTATTTCCCTGGCCAGCATTTTCAGACCTTTCCCATTTGATTT 1301
|||||
DB 150419 AAAATCTGTCATCTTATTTCCCTGGCCAGCATTTTCAGACCTTTCCCATTTGATTT 150360

QY 1302 TTATCGAAGAGGCTCTCAATGTCCACAAAGCATTCAGCATCTTCTTGAGAGCCCGAG 1361
|||||
DB 150359 TTATCGAAGAGGCTCTCAATGTCCACAAAGCATTCAGCATCTTCTTGAGAGCCCGAG 150300

QY 1362 ACACATGATGTTATTCATCAAAACGAAACATCAGGAGATGTAGATGATGAGAAAGAT 1421
|||||
DB 150299 ACACATGATGTTATTCATCAAAACGAAACATCAGGAGATGTAGATGATGAGAAAGAT 150240

QY 1422 TTAGTCATTTTCATGCTTACATTCATATTCATCATATAAG 1462
|||||
DB 150219 TTAGTCATTTTCATGCTTACATTCATATTCATCATATAAG 150199
```

RESULT 15  
AX458379 240 bp DNA linear PAT 08-JUN-2002  
LOCUS AX458379  
DEFINITION Sequence 4 from Patent WO0246411.  
ACCESSION AX458379  
VERSION AX458379.1 GI:21725047  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Lasek, A.W., Krasnow, R. and Baughn, M.R.  
TITLE Intestinal proteins  
JOURNAL Patent: WO 0246411-A 4 13-JUN-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1..240  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 2771041H1"

BASE COUNT 81 a 42 c 44 g 73 t  
ORIGIN

Query Match 14.9% Score 240; DB 6; Length 240;  
Best Local Similarity 100.0%; Pred.No. 1.7e-124;  
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATCCATGCTAAAGGTAACAACTGCACTATATCTGCAATTTATTTGGTATAGCA 60
|||||
DB 1 ATCCATGCTAAAGGTAACAACTGCACTATATCTGCAATTTATTTGGTATAGCA 60

QY 61 GAGGTATGCCAGTACACACTGCTGCTCAGAGAAATTCACACCTTACCTGCGCAG 120
|||||
DB 61 GAGGTATGCCAGTACACACTGCTGCTCAGAGAAATTCACACCTTACCTGCGCAG 120

QY 121 AGACTATGATGAGATGAACTGCTTAATCTAAAGGATCTAATGATGAAATTAAG 180
|||||
DB 121 AGACTATGATGAGATGAACTGCTTAATCTAAAGGATCTAATGATGAAATTAAG 180

QY 181 TATGATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
|||||
DB 181 TATGATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
```

Search completed: November 22, 2002, 03:58:28  
Job time : 4503 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleotide - nucleotide search, using sw model

Run on: November 22, 2002, 01:08:46 ; Search time 332 Seconds  
(without alignments)  
10961.534 Million cell updates/sec

File: us-09-729-454-3

Perfect score: 1616

Sequence: 1 atccatgctaaagtaataa.....atacaaaaaaaaaaaaaa 1616

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum hit seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N\_Geneseq\_101002.\*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	987	61.1	1739	24	ABK33541
2	983	60.8	1138	22	AAF81755
3	880	54.5	1629	24	ABQ86149
4	694	42.9	1137	23	AA576631
5	400	24.8	400	24	ABN61431
6	385	23.9	437	23	AA592089
7	198	12.3	557	24	AB136592
8	192	11.9	275	22	AAF93741
9	182	11.3	817	22	AAK88279

C	10	181	11.2	532	24	ABK27337	Human colon cancer
C	11	93	3.8	583	24	ABJ67799	Human colon tumour
C	12	88	5.4	211	24	ABN60352	Human cancer relat
C	13	88	5.4	611	24	ABN60393	Human cancer relat
C	14	88	5.4	684	24	ABK69141	DNA encoding human
C	15	88	5.4	1177	24	ABK69140	DNA encoding human
C	16	88	5.4	1662	24	ABO86150	DNA encoding human
C	17	88	5.4	1773	24	AA592091	Novel human gene.
C	18	88	5.4	2089	24	ABK69106	DNA encoding novel
C	19	60	3.7	60	24	ABN8360	DNA encoding human
C	20	53	3.3	71	21	AA30578	Human secreted pro
C	21	29	1.8	3676	23	AA588944	Human secreted pro
C	22	24	1.5	580	22	AA498108	DNA encoding novel
C	23	23	1.4	600	22	AA434126	Human colon cancer
C	24	23	1.4	1562	22	AA32540	Mouse P/O-type cal
C	25	23	1.4	3021	22	AA32539	Mouse P/O-type cal
C	26	23	1.4	6542	22	AA32538	Mouse P/O-type cal
C	27	23	1.4	9373	24	ABJ70560	Chemically treated
C	28	23	1.4	9373	24	ABJ70560	Human immune syste
C	29	23	1.4	9373	24	AA561218	Human gene regulat
C	30	23	1.4	9373	24	ABK31317	Signal transductio
C	31	23	1.4	14307	24	ABJ32728	Human immune syste
C	32	22	1.4	243	21	AA64605	Nucleotide sequenc
C	33	22	1.4	345	21	AA25014	Human secreted pro
C	34	22	1.4	396	22	AA594849	Human ovarian can
C	35	22	1.4	396	24	ABT03116	Human ovarian can
C	36	22	1.4	396	24	ABT03116	Ovarian carcinoma
C	37	22	1.4	558	22	ABJ48719	Human prostate exp
C	38	22	1.4	558	22	AAH11526	Human cDNA clone (
C	39	22	1.4	607	21	AAH16283	Human prostate exp
C	40	22	1.4	656	23	AAJ18933	DNA encoding novel
C	41	22	1.4	980	23	AA593510	Plant pathogen ind
C	42	22	1.4	2007	21	AA244859	Human cDNA for nov
C	43	22	1.4	2012	24	ABK34756	Human cDNA sequenc
C	44	22	1.4	2497	22	AAH14473	Atherosclerosis-as
C	45	22	1.4	4158	22	AAK85098	

## ALIGNMENTS

RESULT 1  
ABK33541  
ID ABK33541 standard; cDNA: 1739 BP.  
XX  
XX  
AC ABK33541:  
XX  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA encoding human PRO protein. Seq ID No 11.  
XX  
XX Human; secreted protein. PRO; tumour; lung cancer; colon cancer;  
KW breast cancer; prostate tumour; rectal tumour; liver tumour;  
KW pericyte cell proliferation; chondrocyte cell proliferation;  
XX tumour necrosis factor-alpha; gene; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
XX PN MO200208288-A2.  
XX  
XX PD 31-JAN-2002.  
XX  
XX PF 29-JUN-2001; 2001WO-US21066.  
XX  
XX  
PR 20-JUL-2000; 2000US-219556P.  
PR 25-JUL-2000; 2000US-220585P.  
PR 25-JUL-2000; 2000US-220605P.  
PR 25-JUL-2000; 2000US-220607P.  
PR 25-JUL-2000; 2000US-220624P.  
PR 25-JUL-2000; 2000US-220638P.  
PR 25-JUL-2000; 2000US-220644P.  
PR 25-JUL-2000; 2000US-220666P.  
PR 26-JUL-2000; 2000US-220893P.

```

PR 28-JUL-2000: 2000MO-US20710.
PR 23-AUG-2000: 2000MO-US23522.
PR 24-AUG-2000: 2000MO-US23528.
PR 15-SEP-2000: 2000US-000000P.
PR 10-NOV-2000: 2000MO-US30873.
PR 28-NOV-2000: 2000US-253646P.
PR 01-DEC-2000: 2000MO-US32678.
PR 20-DEC-2000: 2000US-0747259.
PR 20-DEC-2000: 2000MO-US34956.
PR 28-FEB-2001: 2001MO-US06520.
PR 10-MAY-2001: 2001MO-US054280.
PR 25-MAY-2001: 2001MO-US17092.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
XX Grimaldi JC, Gurney AL, Smith Y, Stephan JF, Matanabe CK, Wood WJ;
XX WPI: 2002-172001/22.
XX P-PSDB: AA083597.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
XX useful for treating a PRO related disorder and for diagnosing tumours
XX such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
XX tumour or liver tumour.
XX
XX Claim 2; Figure 11; 359pp; English.
XX
XX The invention relates to one hundred and twenty two nucleic acids
XX encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
XX encode human secreted proteins. The PRO nucleic acids, polypeptides,
XX agonists and antagonists are useful for treating a PRO related disorder.
XX The PRO polypeptides are useful for diagnosing tumours, especially lung
XX cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
XX liver tumour. The PRO polypeptides are useful for stimulating the
XX proliferation of, or gene expression, in pericyte cells, for stimulating
XX the proliferation or differentiation of chondrocyte cells, for
XX stimulating the release of tumour necrosis factor-alpha from human blood,
XX for stimulating or inhibiting the proliferation of normal human dermal
XX fibroblast cells. The PRO polypeptide may also be used as molecular
XX weight markers and for tissue typing. The PRO nucleic acids have
XX applications in molecular biology, including use as hybridisation probes,
XX and in chromosome and gene mapping. ABR33536-ABR33657 represent human
XX PRO protein coding sequences of the invention.
XX
XX Sequence 1739 BP; 545 A; 405 C; 359 G; 430 T; 0 other;
XX
XX Query Match 61.1%; Score 987; DB 24; Length 1739;
XX Best local similarity 100.0%; Pred. No. 0;
XX Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 75 GCACACTGGTGGCTTCAGAGAAATTTCTCAACACCTAGCTCCGAGAGTCTATGTATG 134
XX |||||
XX 3 GCACACTGGTGGCTTCAGAGAAATTTCTCAACACCTAGCTCCGAGAGTCTATGTATG 62
XX
XX 135 GGATTGACAAATCTGTAACTAAGATCCTAATCATGAAATTAAGTATGATTAATTA 194
XX |||||
XX 63 GGATTGACAAATCTGTAACTAAGATCCTAATCATGAAATTAAGTATGATTAATTA 122
XX
XX 195 AGTCACATTTGGCAGCTGTGTTTATTAATGCTCTCGATCAATTTTACAGATTTCCAGA 254
XX |||||
XX 123 AGTCACATTTGGCAGCTGTGTTTATTAATGCTCTCGATCAATTTTACAGATTTCCAGA 182
XX
XX 255 ACTCCACAAAGTTGGTCTGCTTAACCTATCAATCCCTCCGATTATCGAAGAACT 314
XX |||||
XX 183 ACTCCACAAAGTTGGTCTGCTTAACCTATCAATCCCTCCGATTATCGAAGAACT 242
XX
XX 315 CCACAAAGTCTTATTCCTTAAACACACATGATATCAATTAAGCCATAAGAGAGCTG 374
XX |||||
XX 243 CCACAAAGTCTTATTCCTTAAACACACATGATATCAATTAAGCCATAAGAGAGCTG 302
XX
XX 375 AACTCAGATTAAGAAATCATAGAGAACTAGATGAGATCCGACCGAGACCTTTTA 434
XX |||||

```

```

Db 303 AACTCAGATTAAGAAATCATAGAGAACTAGATGAGATCCGACCGAGACCTTTTA 362
Qy 435 CCACGTTGAAACACCCACACACAGGCGACACATATACACAGCCACATCTCTCAACCTTGAG 494
Db 363 CCACGTTGAAACACCCACACACAGGCGACACATATACACAGCCACATCTCTCAACCTTGAG 422
Qy 495 ATACGTTACTGAGGGGAGACACAGCTGACATCTGCTGAGAGGTGAGGGACCACTTGGGAT 554
Db 423 ATACGTTACTGAGGGGAGACACAGCTGACATCTGCTGAGAGGTGAGGGACCACTTGGGAT 482
Qy 555 GCAGAGACCAATATGCGGGGAGATTTCGAGAGGCGAGAGATGCTTCCGACCGCTGATGG 614
Db 483 GCAGAGACCAATATGCGGGGAGATTTCGAGAGGCGAGAGATGCTTCCGACCGCTGATGG 542
Qy 615 CAGGTCTCTCGAGAAAGGTGACTGACTTCAACAAACGCGACCTACCTGCTACGCTCACTC 674
Db 543 CAGGTCTCTCGAGAAAGGTGACTGACTTCAACAAACGCGACCTACCTGCTACGCTCACTC 602
Qy 675 TGTCTGGAGAGGCGAGGTCTCTGTCTGCTGCTGCTCATGCAAGCTGATGAAGAGGTGT 734
Db 603 TGTCTGGAGAGGCGAGGTCTCTGTCTGCTGCTGCTCATGCAAGCTGATGAAGAGGTGT 662
Qy 735 CAGCTCTCTGAGTGCAGAGAACCAAGGCTATGACAGGGTGAATCTTCACTGGCCAGTTTG 794
Db 663 CAGCTCTCTGAGTGCAGAGAACCAAGGCTATGACAGGGTGAATCTTCACTGGCCAGTTTG 722
Qy 795 TCAATGGCAGCTTCCCAAGTCTCACTGTAATGTGGCGCTTCAACCAAAATGCTCAAT 854
Db 723 TCAATGGCAGCTTCCCAAGTCTCACTGTAATGTGGCGCTTCAACCAAAATGCTCAAT 782
Qy 855 TGTGCACTTCTGAGTGCAGAGAACCAAGGCTTCTACTGTGTGAGGCTCAACAGCA 914
Db 783 TGTGCACTTCTGAGTGCAGAGAACCAAGGCTTCTACTGTGTGAGGCTCAACAGCA 842
Qy 915 TCCCGTGTGCGACACCTCACTGATGATTTCTTAAGAACCAAGAAATTTCTTACTTAGA 974
Db 843 TCCCGTGTGCGACACCTCACTGATGATTTCTTAAGAACCAAGAAATTTCTTACTTAGA 902
Qy 975 AACCAAGAAAGGCTCTTTGAAAGTCAATGTGGGTATGAGATTAAGAAATTCGA 1034
Db 903 AACCAAGAAAGGCTCTTTGAAAGTCAATGTGGGTATGAGATTAAGAAATTCGA 962
Qy 1035 ATACATATGATGCTCCAAATGCAACA 1061
Db 963 ATACATATGATGCTCCAAATGCAACA 989

```

RESULT 2  
AAFB1755  
ID AAFB1755 standard; cDNA: 1138 BP.  
XX  
XX AAFB1755;  
XX  
DT 12-JUN-2001 (first entry)  
XX  
DE Human membrane associated protein MEKAP-15 encoding cDNA.  
KW Human; membrane associated protein; MEKAP; diagnosis; cytosolic;  
KW antiinflammatory; anticonvulsant; immunosuppressive; antidiabetic;  
KW antidiabetic; antidiabetic; gene therapy; cell proliferative disorder;  
KW autoimmune disorder; inflammatory disorder; neurological disorder;  
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;  
KW epilepsy; diarrhoea; ss.  
XX  
XX Homo sapiens.  
OS  
PN MO200112662-A2.  
PN 22-FEB-2001.  
XX  
XX 14-AUG-2000: 2000MO-US22315.  
PF  
XX 17-AUG-1999: 90US-0149641.

Filed 14 Aug 2000

```

PR 09-NOV-1999: 99US-0164203.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX 1a1 P, Yue H, Tang YF, Handman O, Burford N, Azimzai Y;
PI Baughn MR, Lu DM, Patterson C;
XX
XX WPI: 2001-16860/17.
DR
DR P-NSDR: AAB/4709.
XX
XX Isolated polypeptide with a human membrane associated protein sequence
PT is useful for the diagnosis, prevention and treatment of cell
PT proliferative, autoimmune/inflammatory, neurological and
PT gastrointestinal disorders.
XX
XX C1a1m 5; (page 159; 173pp; English.
XX
XX A181741 to A181777 encode the human membrane associated proteins
XX (MEMAP) given in AAB/4695 to AAB/4731. MEMAPs have cytosolic,
XX antiinflammatory, anticonvulsant, immunosuppressive, antidiabetic and
XX antiarteriosclerotic activities, which can be used in gene therapy.
XX MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
XX associated with decreased expression of functional MEMAP and antagonists
XX of MEMAP are used to treat a disease or condition associated with
XX overexpression of functional MEMAP. These disorders include cell
XX proliferative, autoimmune/inflammatory, neurological and gastrointestinal
XX disorders. The MEMAP polynucleotides and proteins are also used for the
XX diagnosis of these disorders. Specific examples of these disorders
XX include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
XX MEMAP proteins can be used to screen for compounds which specifically
XX bind MEMAP including antibodies, oligonucleotides, proteins and small
XX molecules. MEMAP polynucleotides can be used to prepare transgenic
XX animals which can be studied to provide information concerning human
XX disease. Anti-MEMAP antibodies are useful in immunoassays for the
XX detection of MEMAP protein and can be used as antagonists to treat or
XX prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
XX can be delivered to target cells with genetic abnormalities with respect
XX to the expression of MEMAP to treat or prevent a disorder associated
XX with MEMAP.
XX
XX Sequence 1138 BP: 314 A; 274 C; 241 G; 289 T; 0 other:
XX
XX Query Match 60.8%; Score 983; DB 22; Length 1138;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 983; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 19 CAAAGTGCACCTTATATCTGCAATTTATTTGTTATACAAAGGATGCCAGTAC 78
XX |
XX 1 CAAAGTGCACCTTATATCTGCAATTTATTTGTTATACAAAGGATGCCAGTAC 60
XX |
XX 79 ACTGTGCTTCAAGAAATTCACACCTAGCTGCCAGAGCTCTATGTATGGAT 138
XX |
XX 61 ACTGTGCTTCAAGAAATTCACACCTAGCTGCCAGAGCTCTATGTATGGAT 120
XX |
XX 139 TGAACAATCTGTAACCTAAAGATCTCATCTGTAAGAAATAGTATATTAATTAAGTC 198
XX |
XX 121 TGAACAATCTGTAACCTAAAGATCTCATCTGTAAGAAATAGTATATTAATTAAGTC 180
XX |
XX 199 ACTATGGCAGCTGTTTATATTAAGCTCTGATCAATTTTACAGTTTCCAGAACTC 258
XX |
XX 181 ACTATGGCAGCTGTTTATATTAAGCTCTGATCAATTTTACAGTTTCCAGAACTC 240
XX |
XX 259 CACAAGCTTTGCTGCTTAAACTTATCATCTCCCTCATTTACTGGAACAACTCCAC 318
XX |
XX 241 CACAAGCTTTGCTGCTTAAACTTATCATCTCCCTCATTTACTGGAACAACTCCAC 300
XX |
XX 319 AATATCTTATTCCTTAAACCTAGTATCATTTAAAGCCCTAAAGAGAGTGAAGT 378
XX |
XX 301 AATATCTTATTCCTTAAACCTAGTATCATTTAAAGCCCTAAAGAGAGTGAAGT 360
XX |
XX 379 CAGATTAAGGAATATCATAGAACTAGATCAGCAGATCCCAAGCCAGCTTACCCA 438
XX |
XX 361 CAGATTAAGGAATATCATAGAACTAGATCAGCAGATCCCAAGCCAGCTTACCCA 420

```

```

OY 439 CGTGAACACCCACAGGCGGCACATAGCAACAGCCACATCTCTAACCCTCGAGATC 498
XX |
XX 421 CGTGAACACCCACAGGCGGCACATAGCAACAGCCACATCTCTAACCCTCGAGATC 480
XX |
OY 499 GTACTGCAAGGGGAGACACAGCTGACATCTGCTGAGAGGTAGGAGACCTTGGAGCAG 558
XX |
XX 481 GTACTGCAAGGGGAGACACAGCTGACATCTGCTGAGAGGTAGGAGACCTTGGAGCAG 540
XX |
OY 559 GAAGCAATATGCGGGGAGATTCCTGAGGGGCAAGATGCTCCCAAGCCCTATATGACG 618
XX |
XX 541 GAAGCAATATGCGGGGAGATTCCTGAGGGGCAAGATGCTCCCAAGCCCTATATGACG 600
XX |
OY 619 TGTCTGAGAAAGTGAAGTACTGCAACAAAGGCACTTACTGAGTCACTCTGT 678
XX |
XX 601 TGTCTGAGAAAGTGAAGTACTGCAACAAAGGCACTTACTGAGTCACTCTGT 660
XX |
OY 679 CTGGAGAGGCGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
XX |
XX 661 CTGGAGAGGCGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
XX |
OY 739 TCTGTGAGTGCAGAGCAACCAAGGCTATGACAGGGTGAATCTTCACTGCGCAGTTGTCAA 798
XX |
XX 721 TCTGTGAGTGCAGAGCAACCAAGGCTATGACAGGGTGAATCTTCACTGCGCAGTTGTCAA 780
XX |
OY 799 TGGCACTTCCCAAGTCCACTGTAAGTGGCTGATCTTAAACCAAAATGCTGAATGTG 858
XX |
XX 781 TGGCACTTCCCAAGTCCACTGTAAGTGGCTGATCTTAAACCAAAATGCTGAATGTG 840
XX |
OY 859 CCAATACCTGGAACACAGAGCAAGAGCTTCTACTGTGTGAGGCTCAACATGCC 918
XX |
XX 841 CCAATACCTGGAACACAGAGCAAGAGCTTCTACTGTGTGAGGCTCAACATGCC 900
XX |
OY 919 CTGTGCGCAGCTACATCATGTATCTTAAGAAAGAAAGTTCTTACTTATAGCAACA 978
XX |
XX 901 CTGTGCGCAGCTACATCATGTATCTTAAGAAAGAAAGTTCTTACTTATAGCAACA 960
XX |
OY 979 AGAAAGAGCTCTTTGAAAGT 1001
XX |
XX 961 AGAAAGAGCTCTTTGAAAGT 983
XX |
XX
XX RESULT 3
XX ID ABO86149
XX ABO86149 standard: DNA; 1629 BP.
XX
XX ABO86149;
XX
XX 10-SEP-2002 (first entry)
XX
XX DE Novel human gene. SEQ ID 20.
XX
XX Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;
XX nocitropic; neuroprotective; immunosuppressive; haemostatic;
XX antiinflammatory; cardiatic; antidiabetic; antihypertensive;
XX cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
XX wound healing disorders; atherosclerosis; Parkinson's disease;
XX Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
XX inflammation; neoplastic disease; nervous system disorder;
XX cardiovascular disorders; pancreatitis; respiratory disorder;
XX hyperproliferation; systemic autoimmune disease; hyper-immunity;
XX developmental abnormality; gastrointestinal ulceration; neuropathy;
XX haematological disease; metabolic disease; sperm dysfunction;
XX thyroid disorder; hypothyroidism; brain damage; colitis;
XX cone photo- transduction deficiency; neurological disease; stroke;
XX angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
XX trachea; thymus; lymph node; muscular system; obesity; anorexia;
XX growth abnormality; precocious puberty; gene; ss.
XX
XX Homo sapiens.
XX
XX W0200250105-A1.
XX

```

```

PD 27-JUN-2002.
XX
XX 17-DEC-2001; 2001WO-US49232.
XX
XX 19-DEC-2000; 2000US-256710P.
XX
XX 20-DEC-2000; 2000US-257048P.
XX
XX 09-JAN-2001; 2001US-260482P.
XX
XX 30-JAN-2001; 2001US-264922P.
XX
XX 06-FEB-2001; 2001US-266797P.
XX
XX 19-MAR-2001; 2001US-276988P.
XX
XX 04-APR-2001; 2001US-281535P.
XX
XX 08-MAY-2001; 2001US-289622P.
XX
XX (SMK ) SMITHKLINE BEECHAM CORP.
XX
XX (SMK ) SMITHKLINE BEECHAM PLC.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Agarwal P, Birkeland M, Cogswell JP, Kohnick KF, Lal Y;
XX Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX WPI: 2002-508784/54.
XX P-PSDB; ABP0984.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
XX or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX Parkinson's disease, Alzheimer's disease, infection, autoimmune
XX disorder -
XX
XX Claim 2(a); Page 235; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
XX which allow it to be secreted extracellularly or membrane associated.
XX The activity of polypeptides of the invention may be described as,
XX cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, nootropic,
XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX cardiant, antituber, virucide, antihypoid, cerebroprotective, anorectic,
XX and metabolic. Polypeptides and polynucleotides of the invention are
XX useful in the treatment, or as a vaccine in the prevention of, cancer,
XX wound healing disorders, infection, atherosclerosis, Parkinson's disease
XX and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX inflammation, neoplastic diseases, nervous system related disorders and
XX cardiovascular disorders, pancreatitis, respiratory disorder,
XX hyperproliferation, systemic autoimmune disease, hyper-immunity,
XX developmental abnormality, gastrointestinal ulceration, neuropathy,
XX haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
XX transduction deficiency, neurological diseases, stroke, angiogenesis,
XX ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
XX trachea, thymus, lymph node and muscular system, obesity, anorexia,
XX growth abnormalities, and alleviation of precocious puberty. The
XX sequences given in records AB086130-AB086184 represent novel human cDNA's
XX of the invention.
XX
XX Sequence 1629 BP; 497 A; 391 C; 329 G; 412 T; 0 other:
XX
XX Query Match 54.5%; Score 880; DB 24; Length 1629;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 182 ATGATAAATATATAGTACATATGACACTGTGTTATATTAGCCTCGATCATTTT 241
XX |||||||
XX 1 ATGATAAATATATAGTACATATGACACTGTGTTATATTAGCCTCGATCATTTT 60
XX
XX 242 ACAGTTTCAGAACCTCCACAAAGCTTGCTGCTCAAACTATACATCCCTCCAT 301
XX |||||||
XX 61 ACAGTTTCAGAACCTCCACAAAGCTTGCTGCTCAAACTATACATCCCTCCAT 120
XX
XX 302 TACTGAAACAACTCCACAAAGTCTTATTCCTAAAGCAGCATGATCATTAAGCA 361
XX |||||||
XX 121 TACTGAAACAACTCCACAAAGTCTTATTCCTAAAGCAGCATGATCATTAAGCA 180
XX
XX 362 CTAACAGAGACTGAATGAGATTAAGAAATCATGAGAACTAGTACGAGATCCCA 421
XX |||||||

```

```

Db 181 CTACAGAGACTGAATGAGATTAAGAAATCATGAGAACTAGTACGAGATCCCA 240
OY 422 CCAGACGCTTCACCCAGCGTGAACACACACACAGCGCCACATAGCAGCGCCAGC 481
Db 241 CCAGACGCTTCACCCAGCGTGAACACACACACAGCGCCACATAGCAGCGCCAGC 400
OY 482 CTCACCCCTCGAGATACGACTGACAGGAGAGACAGCTGCACATCTCTCTGAGGTAG 541
Db 301 CTCACCCCTCGAGATACGACTGACAGGAGAGACAGCTGCACATCTCTCTGAGGTAG 360
OY 542 GACCACTTGGAGCCAGAGAAATATGAGCGGATTTCTGAGGCGCCAGCATCTCTTC 601
Db 361 GACCACTTGGAGCCAGAGAAATATGAGCGGATTTCTGAGGCGCCAGCATCTCTTC 420
OY 602 CCAGCGCTGATGCGAGCTGCTTCAGGAAAGTCACTGATTCACACAGCCATCTAC 641
Db 421 CCAGCGCTGATGCGAGCTGCTTCAGGAAAGTCACTGATTCACACAGCCATCTAC 480
OY 662 GTCAGCTTACTCTGTTCTGAGGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 721
Db 481 GTCAGCTTACTCTGTTCTGAGGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
OY 722 AGTGAAGGGGTGTCAGCTCTCTGAGTGCACAGCAACAGGCTATGACAGGCTATTC 781
Db 541 AGTGAAGGGGTGTCAGCTCTCTGAGTGCACAGCAACAGGCTATGACAGGCTATTC 600
OY 782 ACTGGCACTTGTCTCAATGGCACTTCCCAAGTCCATCTGATGCGCTGATCTTAAC 841
Db 601 ACTGGCACTTGTCTCAATGGCACTTCCCAAGTCCATCTGATGCGCTGATCTTAAC 640
OY 842 ACAATGCTGAATTTGTCAGTCTGAGTGCACAGCAACAGGCTATGACAGGCTATTC 901
Db 661 ACAATGCTGAATTTGTCAGTCTGAGTGCACAGCAACAGGCTATGACAGGCTATTC 720
OY 902 AGCGCTCAACATGCCCTGCTGCTGACATCACTGATGATCTTGAAGCAAGAAAT 961
Db 721 AGCGCTCAACATGCCCTGCTGCTGACATCACTGATGATCTTGAAGCAAGAAAT 780
OY 962 TCTTACTTACGAAACAAAGAAAGACCTCTTGAAGCAAGAAATGAGTAT 1021
Db 781 TCTTACTTACGAAACAAAGAAAGACCTCTTGAAGCAAGAAATGAGTAT 640
OY 1022 ATGAAAAATTCATATCATATAGTGTCTCCAAATGCAACA 1061
Db 841 ATGAAAAATTCATATCATATAGTGTCTCCAAATGCAACA 880
XX
XX RESULT 4
XX AAS76631:
XX ID AAS76631 standard: cDNA: 1137 BP.
XX AC AAS76631:
XX XX
XX XX 13-FEB-2002 (first entry)
XX XX
XX XX DNA encoding novel human diagnostic protein #12435.
XX DE
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN
XX PN W0200175067-A2.
XX PD 11-OCT-2001.
XX XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX XX
XX PA (HYSEQ) HYSEQ INC.
XX XX

```

PI Drmanac RT, Liu C, Tang YT;  
XX WPI: 2001-639362/73.  
DR P-PSDB: ABO12444.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
PS  
XX Claim 1: SEQ ID NO 12435; 103bp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
SQ Sequence 1137 BP; 332 A; 259 C; 254 G; 292 T; 0 other;  
Query Match 42.9%; Score 694; DB 23; Length 1137;  
Best Local Similarity 99.5%; Pred. No. 4,6e-294;  
Matches 944; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 536 GTGAGGACCACTTGGACCCAGAGAAACAAATATGCGCGGATTTCTGAGGCCACGATG 595  
DB 43 GTGAGGACCACTTGGACCCAGAGAAACAAATATGCGCGGATTTCTGAGGCCACGATG 102  
OY 596 TCTTCCCGACCGCTGATGGCAGTGGTCTGAGGAAAGTGACTCTTCAACAACGCGACC 655  
DB 103 TCTTCCCGACCGCTGATGGCAGTGGTCTGAGGAAAGTGACTCTTCAACAACGCGACC 162  
OY 656 TACCTGGTACGTTCACTCTGTTCTGGAGGGCAAGTCTCTGCTGTGCTGATC 715  
DB 163 TACCTGGTACGTTCACTCTGTTCTGGAGGGCAAGTCTCTGCTGTGCTGATC 222  
OY 716 CACCCAGTGAAGGGCTGCTACCTCTGAGTGCAGAGAAACCAAGCTATGACAGGTG 775  
DB 223 CACCCAGTGAAGGGGTGTGACGCTCTGTGAGTGAAGAAACCAAGCTATGACAGGTG 282  
OY 776 ATCTTACTGCGCCAGTTTGTCAATGCGCACTTCCCAAGTCCACTCTGAATGTGCGTATC 835  
DB 283 ATCTTACTGCGCCAGTTTGTCAATGCGCACTTCCCAAGTCCACTCTGAATGTGCGTATC 342  
OY 836 CTAAACCAAAATGCTGAATGTGCGCACTGTCAGACAAGAGCAACCAAGAGGCTTCTAC 895  
DB 343 CTAAACCAAAATGCTGAATGTGCGCACTGTCAGACAAGAGCAACCAAGAGGCTTCTAC 402  
OY 896 TGTGTGAGGCTTCAGACATGCGCTGTGCTGACACTCACTCATGATTTCTAAGAACAG 955  
DB 403 TGTGTGAGGCTTCAGACATGCGCTGTGCTGACACTCACTCATGATTTCTAAGAACAG 462  
OY 956 AAAGTTCTTATCTTACCAAAACAAAGAGGCTCTTTGAAGGTCAATGTGGGTGTA 1015  
DB 463 AAAGTTCTTATCTTACCAAAACAAAGAGGCTCTTTGAAGGTCAATGTGGGTGTA 522  
OY 1016 GAGATTATGAAAAATTAATACATTATGTCTCCAAATGCAACACACTGAAGTCAGTG 1075

DB 523 GAGATTATGAAAAATTAATACATTATGTCTCCAAATGCAACACACTGAAGTCAGTG 582  
OY 1076 GATTCGATGAAATCTGGAATAATTCGAACACAGCTGCTGATTTGATGATGACATC 1135  
DB 583 GATTCGATGAAATCTGGAATAATTCGAACACAGCTGCTGATTTGATGATGACATC 642  
OY 1136 AGATTCAGTGGCAAAAATATTTGATATCCTTGTAGATCAATGACCTATTCAGTCA 1195  
DB 643 AGATTCAGTGGCAAAAATATTTGATATCCTTGTAGATCAATGACCTATTCAGTCA 702  
OY 1196 GAGATGAGTACCTCACCAGGCAATGACAGCACTGAGAGAGAAAAAATACGTCAAT 1255  
DB 703 GAGATGAGTACCTCACCAGGCAATGACAGCACTGAGAGAGAAAAAATACGTCAAT 762  
OY 1256 GTTATTTCCCTGGGCGCAGATTTCAGACCTTTCCCATTTGATGTTTATCCGAAGGCC 1315  
DB 763 GTTATTTCCCTGGGCGCAGATTTCAGACCTTTCCCATTTGATGTTTATCCGAAGGCC 822  
OY 1316 CTCATGTCACAAAGCATTCAGATCTGCTGTGAGAAAGCCAGACATATGTTATC 1375  
DB 823 CTCATGTCACAAAGCATTCAGATCTGCTGTGAGAAAGCCAGACATATGTTATC 882  
OY 1376 ATCAAAACGAAAAACATCAGGAGATGTACAAATGATSCAGAAAGATTAGACTTCAT 1435  
DB 883 ATCAAAACGAAAAACATCAGGAGATGTACAAATGATSCAGAAAGATTAGACTTCAT 942  
OY 1436 GGTATCATTCATATCATCATTAAGAGCAATTTTCAGATGCTAGTG 1484  
DB 943 GGTATCATTCATATCATCATTAAGAGCAATTTTCAGATGCTAGTG 991  
RESULT 5  
ABN61431  
ID ABN61431 standard; cDNA; 400 BP.  
XX  
AC ABN61431;  
XX  
DT 28-JUN-2002 (first entry)  
XX  
DE Human cancer related polynucleotide SEQ ID NO 1398.  
XX  
KW Human; cytosolic; gene expression; gene mapping; tissue profiling;  
KW gene therapy; cancer; tumour; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN M0200214500-A2.  
XX  
PD 21-FEB-2002.  
XX  
PF 16-AUG-2001; 2001WO-0525840.  
XX  
PR 16-AUG-2000; 2000US-226326P.  
XX  
PA (CHIR) CHIRON CORP.  
PA (HSE-) HYSBO INC.  
XX  
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;  
PI Lamson G, Scott EM, Zhang G, Kassam A, Pol D, Labat I;  
XX  
DR WPI: 2002-241905/29.  
XX  
PT New nucleic acid for producing a polypeptide, detecting differentially  
PT expressed genes correlated with a cancerous state of a mammalian cell,  
PT and inhibiting tumor growth -  
XX  
PS Claim 1: SEQ ID NO 1398; 883bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33362)  
CC with cytosolic activity. The polynucleotide is used to produce a  
CC polypeptide, to detect differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell and to inhibit tumour growth. The



AC ABI36592;  
XX  
DT 08-APR-2002 (first entry)  
XX  
DE Human colon tumour antigen polynucleotide SEQ ID NO:181.  
XX  
XX Human: colon cancer; colon tumour antigen; cytostatic; vaccine;  
KM colon tumour metastatic antigen; diagnosis; gene; ss.  
XX  
O5 Homo sapiens.  
XX  
PN W0200196388-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 08-JUN-2001; 2001WO-US18557.  
XX  
PR 09-JUN-2000; 2000US-210899P.  
XX  
PR 20-FEB-2001; 2001US-270216P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
FI Jiang Y, Harlocker SL, Secretist H;  
XX  
DK WPI: 2002-114514/15.  
XX  
PT Novel isolated colon tumor polynucleotide differentially expressed in  
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,  
PT useful for inhibiting development of cancer in patient -  
PS Claim 1; SEQ ID 181; 105pp; English.  
XX  
XX ABI36412 to ABI38645 represent human colon tumour antigen cDNA clones (1)  
CC which were isolated from human colon tumour and colon metastatic tumour  
CC cDNA libraries. (1) have cytostatic activity and can be used in vaccine  
CC production. (1) can be used for stimulating and/or expanding T cells  
CC specific for a tumour protein on contact with the T cells. They are also  
CC useful for inhibiting the development of cancer in a patient. (1) can be  
CC used as probes or primers for nucleic acid hybridisation, for preparing  
CC mutant species primers, or primers for use in genetic constructions. (1)  
CC can be used in the diagnosis of a colon tumour.  
XX  
SQ Sequence 557 BP; 192 A; 112 C; 121 G; 130 T; 2 other;  
XX  
Query Match 12.3%; Score 198; DB 24; Length 557;  
Best Local Similarity 100.0%; Pred. No. 9,6e-77;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 864 ACCTGCACACAGACAGACCAAGGCTTCTACTGTGTAGGCTTCAACACATGCCCTGTG 923  
DB 1 ACCTGCACACAGACAGACCAAGGCTTCTACTGTGTAGGCTTCAACACATGCCCTGTG 60  
OY 924 CTGCACTCACTCACTATGTTCTTAAGACAGAAAGTTCTTATCTAGCAAAAGAAA 983  
DB 61 CTGCACTCACTCACTATGTTCTTAAGACAGAAAGTTCTTATCTAGCAAAAGAAA 120  
OY 944 AGAGCTCTTTGAAAGGTCAAAATGGGTGTAGACATATATGAAAAATTCATACATTA 1043  
DB 121 AGAGCTCTTTGAAAGGTCAAAATGGGTGTAGACATATATGAAAAATTCATACATTA 180  
OY 1044 GTGTCTCCAAATCCAAACA 1061  
DB 181 GTGTCTCCAAATCCAAACA 198  
RESULT 8  
AAK8279  
ID AAK8279 standard; cDNA; 275 BP.  
XX  
XX AAK8279;  
XX  
XX 21-MAY-2001 (first entry)  
XX  
XX

DE cDNA encoding SRT protein isolated from testis tissue SEQ ID 562.  
XX  
XX Human: SRT; gene therapy; gene mapping; tissue typing; ss.  
XX  
O5 Homo sapiens.  
XX  
PN W0200107611-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 21-JUL-2000; 2000WO-US20006.  
XX  
PR 26-JUL-1999; 99US-0145701.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Goddard A, Wood WI;  
XX  
DR WPI: 2001-112729/12.  
XX  
PT New isolated nucleic acid molecule encoding a SRT polypeptide is useful  
PT for production of recombinant SRT polypeptides, gene mapping,  
PT diagnosing genetic disorders and for gene therapy -  
PS Claim 2; Fig 562; 663pp; English.  
XX  
XX Sequences AAK8279 - AAK8279 represent polynucleotide sequences encoding  
CC human SRT proteins. The cDNA sequences are isolated from various  
CC different human tissue cDNA libraries. The invention relates to a method  
CC for detecting cDNA encoding an SRT protein, a vector containing cDNA  
CC encoding SRT, a host cell transformed with the vector, an isolated SRT  
CC polypeptide, and an antibody which binds to SRT. The polynucleotide  
CC sequence can be used in gene therapy and is useful in the recombinant  
CC production of SRT polypeptides, as a hybridisation probe to screen  
CC libraries to isolate cDNAs with sequence identity to SRT polypeptides, to  
CC map the gene encoding the SRT polypeptides and analysing genetic  
CC disorders, tissue typing and disease detection. The SRT  
CC polynucleotide sequences can be used in polymerase chain reaction,  
CC screening for new therapeutic molecules and generation of antisense RNA  
CC and DNA.  
XX  
SQ Sequence 275 BP; 83 A; 65 C; 43 G; 81 T; 3 other;  
XX  
Query Match 11.9%; Score 192; DB 22; Length 275;  
Best Local Similarity 100.0%; Pred. No. 4,2e-74;  
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 81 TGGTGGCTTCAGAAAGAAATTTCTCAACACTAGCTGCCAGAGAGCTATGTGGATTG 140  
DB 14 TGGTGGCTTCAGAAAGAAATTTCTCAACACTAGCTGCCAGAGAGCTATGTGGATTG 73  
OY 141 AACAACTGTAAACTAAAGATCCTATCATGAAATTAAGATGAATTAATTAAGTCA 200  
DB 74 AACAACTGTAAACTAAAGATCCTATCATGAAATTAAGATGAATTAATTAAGTCA 133  
OY 201 TATGGACACTGTTTATATATACCGCTCGATGATTTTATAGTTTCCAGAACTCA 260  
DB 134 TATGGACACTGTTTATATATACCGCTCGATGATTTTATAGTTTCCAGAACTCA 193  
OY 261 CAAGGTTTGGT 272  
DB 194 CAAGGTTTGGT 205  
RESULT 9  
AAK8279  
ID AAK8279 standard; cDNA; 817 BP.  
XX  
XX AAK8279;  
XX  
XX 05-NOV-2001 (first entry)  
XX  
XX Human digestive system antigen coding sequence SEQ ID NO: 595.  
XX  
XX



XX Human, digestive system antigen, gene therapy, cancer, appendicitis;  
KM ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO20015314-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US013324.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232357.  
PR 14-SEP-2000; 2000US-0232358.  
PR 14-SEP-2000; 2000US-0232359.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239355.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-02441826.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249287.  
PR 17-NOV-2000; 2000US-0249289.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.

## RESULT 11

ABL6799/C  
ID ABL6799 standard; cDNA: 583 BP.

AC ABL6799;

DT 08-APR-2002 (first entry)

DE Human colon tumour antigen polynucleotide SEQ ID NO:388.

KM Human; colon cancer; colon tumour antigen; cytosolic; vaccine;

KW colon tumour metastatic antigen; diagnosis; gene; ss.

OS Homo sapiens.

PN W0200196388-A2.

PD 20-DEC-2001.

PF 08-JUN-2001; 2001WO-US18557.

PR 09-JUN-2000; 2000US-210899P.

PR 20-FEB-2001; 2001US-270216P.

PA (CORI-) CORIXA CORP.

PI Jiang Y, Harlocker SL, Secrist H;

DR WPI; 2002-114514/15.

PT Novel isolated colon tumor polynucleotide differentially expressed in  
colon tumor or colon metastatic tumor and polypeptides encoded by them,  
useful for inhibiting development of cancer in patient -

PS Claim 1; SEQ ID 388; 105pp; English.

CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (1)  
CC which were isolated from human colon tumour and colon metastatic tumour  
CC cDNA libraries. (1) have cytosolic activity and can be used in vaccine  
CC production. (1) can be used for stimulating and/or expanding T cells  
CC specific for a tumour protein on contact with the T cells. They are also  
CC useful for inhibiting the development of cancer in a patient. (1) can be  
CC used as probes or primers for nucleic acid hybridisation, for preparing  
CC mutant species primers, or primers for use in genetic constructions. (1)  
CC can be used in the diagnosis of a colon tumour.

3Q Sequence 583 BP; 186 A; 110 C; 126 G; 157 T; 4 other;

Query Match 5.8%; Score 93; DB 24; Length 583;

Best Local Similarity 99.3%; Pred. No. 9.7e-31;

Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1062 CACTGAAGTCAGTCGATCTGCATCATCTGGAATTCGACACACGCTTGCTGCGATT 1121

Y 144 CACTGAAGTCAGTCGATCTGCATCAATCTGGAATTCGACACACGCTTGCTGCGATT 85

Y 1122 TGGATAGGAACATCAATCCATCCAGTGGCAAAATATGTTATCCTTGATAGATCAATGA 1181

Y 84 TGGATAGGAACATCAATCCATCCAGTGGCAAAATATGTTATCCTTGATAGATCAATGA 25

Y 1182 CCTATTCAGTCAAGAGATGAGT 1205

Y 24 CCTATTCAGTCAAGAGATGAGT 1

## RESULT 12

ABN60352  
ID ABN60352 standard; cDNA: 211 BP.

KC ABN60352;

Y 28-JUN-2002 (first entry)

XX Human cancer related polynucleotide SEQ ID NO 319.

KM Human; cytosolic; gene expression; gene mapping; tissue profiling;

KW gene therapy; cancer; tumour; gene; ss.

OS Homo sapiens.

PN W0200214500-A2.

PD 21-FEB-2002.

PF 16-AUG-2001; 2001WO-US25840.

PR 16-AUG-2000; 2000US-226326P.

PA (CHIR-) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Kandarzo F;

PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;

DR WPI; 2002-241905/29.

PT New nucleic acid for producing a polypeptide, detecting differentially  
expressed genes correlated with a cancerous state of a mammalian cell,  
and inhibiting tumor growth -

PS Claim 1; SEQ ID NO 319; 883pp + Sequence Listing; English.

CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)  
CC with cytostatic activity. The polynucleotide is used to produce a  
CC polypeptide, to detect differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell and to inhibit tumour growth. The  
CC polynucleotide is used as a probe in mapping and tissue profiling. The  
CC encoded polypeptide and antibodies to the polypeptide can also be used  
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for  
CC gene therapy.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WPI  
CC at ftp.wipo.int/pub/published\_pat\_sequences.

3Q Sequence 211 BP; 75 A; 68 C; 32 G; 36 T; 0 other;

Query Match 5.4%; Score 88; DB 24; Length 211;

Best Local Similarity 100.0%; Pred. No. 1.6e-28;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 CATTAAGCCACTAAGACAGAGACTGAGTGAATTAAGAAATCATAGAGAACTAGATC 419

DB 31 CATTAAGCCACTAAGACAGAGACTGAGTGAATTAAGAAATCATAGAGAACTAGATC 90

QY 411 AGCAGATCCGACCCGACCTTTCACCA 438

DB 91 AGCAGATCCGACCCGACCTTTCACCA 118

## RESULT 13

ABN60393  
ID ABN60393 standard; cDNA: 611 BP.

AC ABN60393;

Y 28-JUN-2002 (first entry)

DE Human cancer related polynucleotide SEQ ID NO 360.

KM Human; cytosolic; gene expression; gene mapping; tissue profiling;

KW gene therapy; cancer; tumour; gene; ss.

OS Homo sapiens.

PN W0200214500-A2.

```

XX 21-FEB-2002.
PD
XX
XX 16-AUG-2001; 2001MO-US25840.
PF
XX
PR 16-AUG-2000; 2000US-226326P.
XX
XX (CHIR) CHIRON CORP.
PA (HYSH-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PD Lamsou G, Scott EM, Zhang G, Kassam A, Pol D, Labal I;
XX WPI: 2002-241905/29.
DR
XX
XX New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.
XX
XX Claim 1: SEQ ID NO 360; 883bp + Sequence Listing: English.
XX
XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published\_pat\_sequences.
XX
SU Sequence 611 BP; 188 A; 163 C; 119 G; 141 T; 0 other;
Query Match 5.4%; Score 88; DB 24; Length 611;
Best Local Similarity 100.0%; Pred. No. 1.5e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dy 351 CATTAAGCCCTACAGAGACTGACATTAAGCAATCTACAGAACTGATC 410
Db 305 CATTAAGCCCTACAGAGACTGACATTAAGCAATCTACAGAACTGATC 364
Dy 411 AGCAGATCCACCCAGACCTTTACCCCA 438
Db 365 ACCAGATCCACCCAGACCTTTACCCCA 392

```

## RESULT 14

```

ABK69141
ID ABK69141 standard; cDNA; 684 BP.
AC
XX
AC ABK69141;
XX
DT 02-JUL-2002 (first entry)
XX
DE DNA encoding human secreted protein, SEQ ID NO 65.
XX
XX Human; secreted protein; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; breast; liver; ischemia;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW angio genesis; nervous system disorder; Alzheimer's disease; infection;
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;
KW epithelial cell proliferation; organ transplantation; food additive;
KW food storage; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200224721-A1.
PN
XX
XX 28-MAR-2002.
PD
XX
XX 09-JAN-2001; 2001MO-US00544.
PF

```

```

XX 20-SEP-2000; 2000US-234211P.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Komatsoulis GA, Baker KP, Rosen CH, Birse CE, Soppet DR, Olsen HS;
PI Kl J, Fiscella M, Moore PA, Wei P, Edner R, Duan DR, Shi Y;
XX Choi GH;
XX
XX WPI: 2002-330012/36.
DR
DR P-PDB: AAU96220.
XX
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Claim 1: Page 491; 562pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) encoding a
CC human secreted protein (II). (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to (II) can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischemia, angio genesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal
CC infection. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. ABK69078-ABK69143
CC represent human secreted protein coding sequences, PCR primers and
CC related sequences used in cloning and expression of the secreted
CC proteins described in examples of the invention.
XX
SU Sequence 684 BP; 189 A; 185 C; 147 G; 158 T; 5 other;
Query Match 5.4%; Score 88; DB 24; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.5e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dy 351 CATTAAGCCCTACAGAGACTGACATTAAGCAATCTACAGAACTGATC 410
Db 294 CATTAAGCCCTACAGAGACTGACATTAAGCAATCTACAGAACTGATC 353
Dy 411 AGCAGATCCACCCAGACCTTTACCCCA 438
Db 354 AGCAGATCCACCCAGACCTTTACCCCA 381

```

## RESULT 15

```

ABK69140
ID ABK69140 standard; cDNA; 1177 BP.
AC
XX
AC ABK69140;
XX
DT 02-JUL-2002 (first entry)
XX
DE DNA encoding human secreted protein, SEQ ID NO 64.
XX
XX Human; secreted protein; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; breast; liver; ischemia;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW angio genesis; nervous system disorder; Alzheimer's disease; infection;
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;
KW epithelial cell proliferation; organ transplantation; food additive;
KW food storage; gene; ss.

```

```

XX OS Homo sapiens.
XX PN WO200224721-A1.
XX PD 28-MAR-2002.
XX PF 09-JAN-2001; 2001WO-US00544.
XX PR 20-SEP-2000; 2000US-234211P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Komatsoulis GA, Baker KP, Rosen CA, Birse CE, Soppet DR, Olsen HS,
XX PI NJ, Fliscella M, Moore PA, Wel P, Ebner R, Duan DR, Shi Y;
XX PI Chol GH;
XX DR WPI: 2002-330012/36.
XX DR P-PSDB; AA096219.
XX PT Isolated nucleic acid molecule encoding a human secreted protein is
XX PS used in preventing, treating or ameliorating a medical condition -
XX PS Claim 1; Page 490; 562pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) encoding a
XX CC human secreted protein (II). (I) and (II) are used to prevent, treat or
XX CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
XX CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
XX CC diagnosing a pathological condition or susceptibility to a pathological
XX CC condition. The antibodies to (II) can also be used in alleviating
XX CC symptoms associated with the disorders and in diagnostic immunoassays
XX CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
XX CC Disorders which are diagnosed or treated include autoimmune diseases
XX CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
XX CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
XX CC cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous
XX CC system disorders e.g. Alzheimer's disease, infections caused by
XX CC bacteria, viruses and fungi and ocular disorders e.g. corneal
XX CC infection. The polypeptides can also be used to aid wound healing and
XX CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
XX CC maintain organs before transplantation, for supporting cell culture of
XX CC primary tissues, to regenerate tissues and in chemotaxis. The
XX CC polypeptides can also be used as a food additive or preservative to
XX CC increase or decrease storage capabilities. ABK69078-ABK69143
XX CC represent human secreted protein coding sequences, PCR primers and
XX CC related sequences used in cloning and expression of the secreted
XX CC proteins described in examples of the invention.
XX SO Sequence 1177 BP; 347 A; 289 C; 253 G; 273 T; 15 other:

Query Match 5.4%; Score 88; DB 24; Length 1177;
Best Local Similarity 100.0%; Pred. No. 1.4e-28;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 CATTAAAGCCACTAACAGAGACTGAGTCAATTAAGAAATCATAGAGAACTAGATC 410
   |||||||
Db 357 CATTAAAGCCACTAACAGAGACTGAGTCAATTAAGAAATCATAGAGAACTAGATC 416
QY 411 AGCAGATCCAGCCGACGACCTTCACCA 438
   |||||||
Db 417 AGCAGATCCAGCCGACGACCTTCACCA 444

```

Search completed: November 22, 2002, 02:43:20  
 Job time : 340 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2002, 02:35:41 ; Search time 83 Seconds  
(Without alignments)  
5970.960 Million cell updates/sec

Title: US-09-729-454-3

Perfect score: 1616  
Sequence: 1 atccatgctaaagtaaac.....aatacaaaaaaaaaaaaaa 1616

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 44362 seqs, 15338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database: Issued Patents, NA:\*

1: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/lna/PCRNUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/lna/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
C 1	22	1	4	607	4 US-09-385-982-288
C 2	20	1	2	1967	4 US-08-398-393A-23
C 3	20	1	2	1968	4 US-08-937-540-3
C 4	20	1	2	29604	3 US-08-781-891-207
C 5	19	1	2	259	3 US-09-103-359-13
C 6	19	1	2	425	4 US-09-247-155-73
C 7	19	1	2	485	4 US-09-385-982-31
C 8	19	1	2	1005	3 US-09-103-359-4
C 9	19	1	2	1356	1 US-08-426-169-2
C 10	19	1	2	1356	4 US-09-233-813-2
C 11	19	1	2	1356	5 PCT-US95-09470-2
C 12	19	1	2	1615	4 US-08-308-814-1
C 13	19	1	2	1809	1 US-08-455-001-1
C 14	19	1	2	1809	5 PCT-US95-11869-1
C 15	19	1	2	2038	4 US-09-276-531-37
C 16	19	1	2	2304	1 US-08-464-266-1
C 17	19	1	2	2304	1 US-08-464-272-1
C 18	19	1	2	2304	4 US-08-464-514-1
C 19	19	1	2	2304	4 US-08-486-403-1
C 20	19	1	2	2981	4 US-08-257-073-2
C 21	19	1	2	2981	2 US-08-184-009-119
C 22	19	1	2	2981	2 US-08-458-356-119
C 23	19	1	2	2981	4 US-08-460-736-119
C 24	19	1	2	3107	4 US-08-213-419B-1
C 25	19	1	2	6124	4 US-08-213-419B-3
C 26	19	1	1	45	4 US-09-025-639-1
C 27	18	1	1	45	4 US-09-333-237-1

C 28	18	1	1	60	4 US-09-410-960-1	Sequence 1, App1
C 29	18	1	1	60	4 US-09-410-960-3	Sequence 3, App1
C 30	18	1	1	60	4 US-09-410-960-4	Sequence 4, App1
C 31	18	1	1	269	4 US-09-149-476-142	Sequence 142, App
C 32	18	1	1	567	4 US-09-385-982-427	Sequence 427, App
C 33	18	1	1	750	4 US-09-352-990-3	Sequence 3, App1
C 34	18	1	1	809	4 US-08-441-629-1	Sequence 1, App1
C 35	18	1	1	809	3 US-08-776-207-1	Sequence 1, App1
C 36	18	1	1	809	5 PCT-US95-09172-1	Sequence 1, App1
C 37	18	1	1	815	4 US-09-288-143-24	Sequence 24, App1
C 38	18	1	1	897	4 US-08-899-330-14	Sequence 14, App1
C 39	18	1	1	904	4 US-09-171-209-24	Sequence 24, App1
C 40	18	1	1	1356	1 US-07-603-133B-1	Sequence 1, App1
C 41	18	1	1	1356	1 US-08-426-169-3	Sequence 3, App1
C 42	18	1	1	1356	5 PCT-US95-09470-3	Sequence 3, App1
C 43	18	1	1	1460	4 US-09-370-838-80	Sequence 80, App1
C 44	18	1	1	1460	4 US-09-370-838-80	Sequence 80, App1
C 45	18	1	1	1460	4 US-09-370-838-80	Sequence 80, App1

#### ALIGNMENTS

```
RESULT 1
US-09-385-982-288/c
; Sequence 288, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDSE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CODNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; EARLIER FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 288
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(607)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-288
Query Match 1.4%; Score 22; DB 4; Length 607;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1595 TAAATACACAAAAA 1616
DB 35 TAAATACACAAAAA 14
RESULT 2
US-09-398-395A-23
; Sequence 23, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
```

```

; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatsSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 1967
; TYPE: DNA
; ORGANISM: Salvia officinalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)...(1785)
; OTHER INFORMATION: 1,8-cineole synthase
US-09-398-3954-23

```

```

Query Match          1.2%: Score 20; DB 4; Length 1967;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1597 AATACAAAAAAAAAAAAA 1616
Db 1939 AATACAAAAAAAAAAAAA 1958

```

```

RESULT 3
US-08-937-540-3
; Sequence 3, Application US/08937540
; Patent No. 5891697
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Wise, Mitchell L
; APPLICANT: Savage, Thomas J
; APPLICANT: Katalifa, Eva J
; TITLE OF INVENTION: Monoterpene Synthases from Common Sage
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS
; STREET: 1420 FIFTH AVENUE
; CITY: SEATTLE
; STATE: WASHINGTON
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC Compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,540
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR11254
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206 695 1718
; TELEFAX: 206 224 0779
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1968 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

```

```

; ORGANISM: Salvia officinalis
; IMMEDIATE SOURCE:
; CLONE: 1,8-cineole synthase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14..1788
US-08-937-540-3

```

```

Query Match          1.2%: Score 20; DB 2; Length 1968;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1597 AATACAAAAAAAAAAAAA 1616
Db 1940 AATACAAAAAAAAAAAAA 1959

```

```

RESULT 4
US-08-781-891-207
; Sequence 207, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC Compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620lenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-207

```

```

Query Match          1.2%: Score 20; DB 3; Length 29604;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1597 AATACAAAAAAAAAAAAA 1616
Db 12364 AATACAAAAAAAAAAAAA 12383

```

```

RESULT 5
US-09-103-359-13/c
; Sequence 13, Application US/09103359

```

```

Patent No. 6057108
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Guebler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Yuc, Henry
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN ARF-RELATED PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/103,359
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0537 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-855-0572
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GBLANOT01
CLONE: 2614213H1
US-09-103-359-13
Query Match
Best Local Similarity 1.28; Score 19; DB 3; Length 259;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 939 TGTATCTAAGAACAAAGAA 957
DB 34 TGTATCTAAGAACAAAGAA 16

```

# RESULT 6

```

US-09-247-155-73
Sequence 73, Application US/09247155A
Patent No. 6312922
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclerly, Aymeric
APPLICANT: Bouquelieret, Lydie
TITLE OF INVENTION: Complementary DNAs
FILE REFERENCE: GENSRT.021A
CURRENT APPLICATION NUMBER: US/09/247,155A
EARLIER FILING DATE: 1999-02-09
EARLIER APPLICATION NUMBER: 60/074,121
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/081,563
EARLIER FILING DATE: 1998-04-13

```

```

EARLIER APPLICATION NUMBER: 60/096,116
EARLIER FILING DATE: 1998-08-10
EARLIER APPLICATION NUMBER: 60/099,273
EARLIER FILING DATE: 1998-10-04
NUMBER OF SEQ ID NOS: 182
SOFTWARE: Patent.pm
SEQ ID NO 73
LENGTH: 425
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 55..291
FEATURE:
NAME/KEY: sig-peptide
LOCATION: 55..255
OTHER INFORMATION: Von Heljne matrix
OTHER INFORMATION: score 4.4
FEATURE:
NAME/KEY: polyA-signal
LOCATION: 390..395
FEATURE:
NAME/KEY: polyA-site
LOCATION: 410..425
US-09-247-155-73

```

```

Query Match
Best Local Similarity 1.28; Score 19; DB 4; Length 425;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1598 ATACACAAAAAATAAAAA 1616
DB 405 ATACACAAAAAATAAAAA 423

```

# RESULT 7

```

US-09-385-982-31/C
Sequence 31, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 485
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(485)
OTHER INFORMATION: n - A,T,C or G
US-09-385-982-31

```

```

Query Match
Best Local Similarity 1.28; Score 19; DB 4; Length 485;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 939 TGTATCTAAGAACAAAGAA 957
DB 54 TGTATCTAAGAACAAAGAA 36

```





Host Local Similarity 100.0%; Pred. No. 28;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TGATGAGAAAGATTAGT 1426  
|||||  
DB 704 TGATGAGAAAGATTAGT 722

## RESULT 11

PCT-US95-09470-2

Sequence 2, Application PC/TUS9509470

GENERAL INFORMATION:

APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL CENTER

TITLE OF INVENTION: DNA VACCINES AGAINST ROTAVIRUS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: Fish &amp; Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/09470

FILING DATE: 26-JUL-95

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/426,169

FILING DATE: 20-APR-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: FASSO, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 04020/049W01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)542-5070

TELEFAX: (617) 542-8906

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1356 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

ORGANISM: Human rotavirus VP6

PCT-US95-09470-2

Query Match

Host Local Similarity 100.0%; Pred. No. 28;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TGATGAGAAAGATTAGT 1426

|||||

DB 704 TGATGAGAAAGATTAGT 722

RESULT 12

US-08-308-814-1

Sequence 1, Application US/08308814

Patent No. 6268476

GENERAL INFORMATION:

APPLICANT: Flanagan, John G.

TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE &amp; COCKFIELD

STREET: 60 State Street

CITY: Boston

CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/308,814  
FILING DATE: 19-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMT-011  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1615 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..636  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: 10..69  
US-08-308-814-1

Query Match 1.2%; Score 19; DB 4; Length 1615;  
Host Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1598 ATACCAAAAAAAAAAAAAA 1616  
|||||  
DB 1534 ATACCAAAAAAAAAAAAAA 1532

RESULT 13  
US-08-455-001-1  
Sequence 1, Application US/08455001  
Patent No. 5795734  
GENERAL INFORMATION:  
APPLICANT: Flanagan, John G.  
TITLE OF INVENTION: EPH Receptor Ligands, and Uses Related  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII (text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,001  
FILING DATE: 31 MAY 1995  
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMI-011CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1809 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 265..891  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..264  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 892..1809  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: 255..324  
US-08-455-001-1

Query Match 1.2% Score 19; DB 1; Length 1809;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

27 1598 ATACACAAAAA 1616  
Db 1789 ATACACAAAAA 1807

RESULT 14  
PCT-US95-11869-1  
Sequence 1, Application PC/TUS9511869  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: EPH Receptor ligands, and Uses Related  
TITLE OF INVENTION: Thetelo  
NUMBER OF SEQUENCES: 5  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE:  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/11869  
FILING DATE: 19-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMI-011CP2C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1809 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 265..891  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..264

FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 892..1809  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: 255..324  
PCT-US95-11869-1

Query Match 1.2% Score 19; DB 5; Length 1809;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1598 ATACACAAAAA 1616  
Db 1789 ATACACAAAAA 1807

RESULT 15  
US-09-276-531-37/C  
Sequence 37, Application US/09276531  
Patent No. 6183968  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Reddy, Koopa  
APPLICANT: Guegler, Karl J.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/276,531  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/079,677  
FILING DATE: March 27, 1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lynn E. Murry, Ph.D.  
REGISTRATION NUMBER: 42,918  
REFERENCE/DOCKET NUMBER: PA-0008 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 845-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2038 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSTUTOS  
CLONE: 838871  
US-09-276-531-37

Query Match 1.2% Score 19; DB 4; Length 2038;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 353 TTAAAGCCACTAACAGAGA 371  
|||||  
Db 1071 TTAAAGCCACTAACAGAGA 1053

Search completed: November 22, 2002, 04:31:19  
Job time : 128 secs

US 097294540DP1



Creation date: 12-01-2003  
Indexing Officer: IMULAMBA - ISIDORE MULAMBA  
Team: OIPEBackFileIndexing  
Dossier: 09729454

Legal Date: 11-27-2002

No.	Doccode	Number of pages
1	CTNF	9
2	892	1
3	NPL	3
4	NPL	8
5	NPL	2
6	NPL	4

Total number of pages: 27

Remarks:

Order of re-scan issued on .....